USE OF PROTEINS AND PEPTIDES ENCODED BY THE GENOME OF A NOVEL SARS-ASSOCIATED CORONAVIRUS STRAIN

The present invention relates to a novel strain of severe acute respiratory syndrome (SARS)-associated coronavirus derived from a sample recorded under No. 031589 and collected in Hanoi (Vietnam), to nucleic acid molecules derived from its genome, to the proteins and peptides encoded by said nucleic acid molecules and to their applications, in particular as diagnostic reagents and/or as vaccine.

Coronavirus is a virus containing single-stranded RNA, of positive polarity, of approximately 30 kilobases which replicates in the cytoplasm of the host cells; the 5′ end of the genome has a capped structure and the 3′ end contains a polyA tail. This virus is enveloped and comprises, at its surface, peplomeric structures called spicules.

The genome comprises the following open reading frames or ORFs, from its 5′ end to its 3′ end: ORF1a and ORF1b corresponding to the proteins of the transcription-replication complex, and ORF-S, ORF-E, ORF-M and ORF-N corresponding to the structural proteins S, E, M and N. It also comprises ORFs corresponding to proteins of unknown function encoded by: the region situated between ORF-S and ORF-E and overlapping the latter, the region situated between ORF-M and ORF-N, and the region included in ORF-N.

The S protein is a membrane glycoprotein (200-220 kDa) which exists in the form of spicules or spikes emerging from the surface of the viral envelope. It is responsible for the attachment of the virus to the receptors of the host cell and for inducing the fusion of the viral envelope with the cell membrane.

The small envelope protein (E), also called sM (*small membrane*), which is a nonglycosylated transmembrane protein of about 10 kDa, is the protein present in the smallest quantity in the virion. It plays a powerful role in the coronavirus budding process which occurs at the level of the intermediate compartment in the endoplasmic reticulum and the Golgi apparatus.

The M protein or matrix protein (25-30 kDa) is a more abundant membrane glycoprotein which is integrated into the viral particle by an M/E interaction, whereas the incorporation of S into the particles is directed by an S/M interaction. It appears to be important for the viral maturation of coronaviruses and for the determination of the site where the viral particles are assembled.

The N protein or nucleocapsid protein (45-50 kDa) which is the most conserved among the coronavirus structural proteins is necessary for encapsidating the genomic RNA and then for directing its incorporation into the virion. This protein is probably also involved in the replication of the RNA.

When the host cell is infected, the reading frame (ORF) situated in 5′ of the viral genome is translated into a polyprotein which is cleaved by the viral proteases and then releases several nonstructural proteins such as the RNA-dependent RNA polymerase (Rep) and the ATPase helicase (Hel). These two proteins are involved in the replication of the viral genome and in the generation of transcripts which are used in the synthesis of the viral proteins. The mechanisms by which these subgenomic mRNAs are produced are not completely understood; however, recent facts indicate that the sequences for regulation of transcription at the 5′ end of each gene represent signals which regulate the discontinuous transcription of the subgenomic mRNAs.

The proteins of the viral membrane (S, E and M proteins) are inserted into the intermediate compartment, whereas the replicated RNA (+ strand) is assembled with the N (nucleocapsid) protein. This protein-RNA complex then combines with the M protein contained in the membranes of the endoplasmic reticulum and the viral particles form when the nucleocapsid complex buds into the endoplasmic reticulum. The virus then migrates across the Golgi complex and eventually leaves the cell, for example by exocytosis. The site of attachment of the virus to the host cell is at the level of the S protein.

Coronaviruses are responsible for 15 to 30% of colds in humans and for respiratory and digestive infections in animals, especially cats (FIPV: *Feline infectious peritonitis virus*), poultry (IBV: *Avian infectious bronchitis virus*), mice (MHV: *Mouse hepatitis virus*), pigs (TGEV: *Transmissible gastroenteritis virus*, PEDV: *Porcine Epidemic diarrhea virus*, PRCoV: *Porcine Respiratory Coronavirus*, HEV: *Hemagglutinating encephalomyelitis Virus*) and bovines (BCoV: *Bovine coronavirus*).

In general, each coronavirus affects only one species; in immunocompetent individuals, the infection induces optionally neutralizing antibodies and cell immunity, capable of destroying the infected cells.

An epidemy of atypical pneumonia, called severe acute respiratory syndrome (SARS) has spread in various countries (Vietnam, Hong Kong, Singapore, Thailand and Canada) during the first quarter of 2003, from an initial focus which appeared in China in the last quarter of 2002. The severity of this disease is such that its mortality rate is about 3 to 6%. The determination of the causative agent of this disease is underway by numerous laboratories worldwide.

In March 2003, a new coronavirus (SARS-CoV or SARS virus) was isolated, in association with cases of severe acute respiratory syndrome (T.G. KSIAZEK et al., The New England Journal of Medicine, 2003, 348, 1319-1330; C. DROSTEN et al., The New England Journal of Medicine, 2003, 348, 1967-1976; Peiris et al., Lancet, 2003, 361, 1319).

Genomic sequences of this new coronavirus have thus been obtained, in particular those of the Urbani isolate (Genbank accession No. AY274119.3 and A. MARRA et al., Science, May 1, 2003, 300, 1399-1404) and the Toronto isolate (Tor2, Genbank accession No. AY278741 and A. ROTA et al., Science, 2003, 300, 1394-1399).

The organization of the genome is comparable with that of other known coronaviruses, thus making it possible to confirm that SARS-CoV belongs to the *Coronaviridae* family; open reading frames ORF1a and 1b and open reading frames corresponding to the S, E, M and N proteins, and to proteins encoded by: the region situated between ORF-S and ORF-E (ORF3), the region situated between ORF-S and ORF-E and overlapping ORF-E (ORF4), the region situated between ORF-M and ORF-N (ORF7 to ORF11) and the region corresponding to ORF-N (ORF13 and ORF14), have in particular been identified.

Seven differences have been identified between the sequences of the Tor2 and Urbani isolates; 3 correspond to silent mutations (c/t at position 16622 and a/g at position 19064 of ORF1b, t/c at position 24872 of ORF-S) and 4 modify the amino acid sequence of respectively: the proteins encoded by ORF1a (c/t at position 7919 corresponding to the A/V mutation), the S protein (g/t at position 23220 corresponding to the A/S mutation), the protein encoded by ORF3 (a/g at position 25298 corresponding to the R/G mutation) and the M

protein (t/c at position 26857 corresponding to the S/P mutation).

In addition, phylogenetic analysis shows that SARS-CoV is distant from other coronaviruses and that it did not appear by mutation of human respiratory coronaviruses nor by recombination between known coronaviruses (for a review, see Holmes, J.C.I., 2003, 111, 1605-1609).

The determination and the taking into account of new variants are important for the development of reagents for the detection and diagnosis of SARS which are sufficiently sensitive and specific, and immunogenic compositions capable of protecting populations against epidemics of SARS.

The inventors have now identified another strain of SARS-associated coronavirus which is distinguishable from the Tor2 and Urbani isolates.

The subject of the present invention is therefore an isolated or purified strain of severe acute respiratory syndrome-associated human coronavirus, characterized in that its genome has, in the form of complementary DNA, a serine codon at position 23220-23222 of the gene for the S protein or a glycine codon at position 25298-25300 of the gene for ORF3, and an alanine codon at position 7918-7920 of ORF1a or a serine codon at position 26857-26859 of the gene for the M protein, said positions being indicated in terms of reference to the Genbank sequence AY274119.3.

According to an advantageous embodiment of said strain, the DNA equivalent of its genome has a sequence corresponding to the sequence SEQ ID No: 1; this coronavirus strain is derived from the sample collected from the bronchoaleveolar washings from a patient suffering from SARS, recorded under the No. 031589 and collected at the Hanoi (Vietnam) French hospital.

In accordance with the invention, said sequence SEQ ID No: 1 is that of the deoxyribonucleic acid corresponding to the ribonucleic acid molecule of the genome of the isolated coronavirus strain as defined above.

The sequence SEQ ID No: 1 is distinguishable from the Genbank sequence AY274119.3 (Tor2 isolate) in that it possesses the following mutations:

- g/t at position 23220; the alanine codon (gct) at position 577 of the amino acid sequence of the Tor2 S protein is replaced by a serine codon (tct),
- a/g at position 25298; the arginine codon (aga) at position 11 of the amino acid sequence of the protein encoded by the Tor2 ORF3 is replaced by a glycine codon (gga).

In addition, the sequence SEQ ID No: 1 is distinguishable from the Genbank sequence AY278741 (Urbani isolate) in that it possesses the following mutations:

- t/c at position 7919; the valine codon (gtt) in position 2552 of the amino acid sequence of the protein encoded by ORF1a is replaced by an alanine codon (gct),
- t/c at position 16622: this mutation does not modify the amino acid sequence of the proteins encoded by ORF1b (silent mutation),
- g/a at position 19064: this mutation does not modify the amino acid sequence of the proteins encoded by ORF1b (silent mutation),
- c/t at position 24872: this mutation does not modify the amino acid sequence of the S protein, and
- c/t at position 26857: the proline codon (ccc) at position 154 of the amino acid sequence of the M protein is replaced by a serine codon (tcc).

Unless otherwise stated, the positions of the nucleotide and peptide sequences are indicated with reference to the Genbank sequence AY274119.3.

The subject of the present invention is also an isolated or purified polynucleotide, characterized in that its sequence is that of the genome of the isolated coronavirus strain as defined above.

According to an advantageous embodiment of said polynucleotide, it has the sequence SEQ ID No: 1.

The subject of the present invention is also an isolated or purified polynucleotide, characterized in that its sequence hybridizes under high stringency conditions with the sequence of the polynucleotide as defined above.

The terms "isolated or purified" mean modified "by the hand of humans" from the natural state; in other words if an object exists in nature, it is said to be isolated or purified if it is modified or extracted from its natural environment or both. For example, a polynucleotide or a protein/peptide naturally present in a living organism is neither isolated nor purified; on the other hand, the same polynucleotide or protein/peptide separated from coexisting molecules in its natural environment, obtained by cloning, amplification and/or chemical synthesis is isolated for the purposes of the present invention. Furthermore, a polynucleotide or a protein/peptide which is introduced into an organism by transformation, genetic manipulation or by any other method, is "isolated" even if it is present in said organism. The term purified as used in the present invention means that the proteins/peptides according to the invention are essentially free of association with the other proteins or polypeptides, as is for example the product purified

from the culture of recombinant host cells or the product purified from a nonrecombinant source.

For the purposes of the present invention, high stringency hybridization conditions are understood to mean temperature and ionic strength conditions chosen such that they make it possible to maintain the specific and selective hybridization between complementary polynucleotides.

By way of illustration, high stringency conditions for the purposes of defining the above polynucleotides are advantageously the following: the DNA-DNA or DNA-RNA hybridization is performed in two steps: (1) prehybridization at 42°C for 3 hours in phosphate buffer (20 mM, pH 7.5) containing 5 × SSC (1 × SSC corresponds to a 0.15 M NaCl + 0.015 M sodium citrate solution), 50% formamide, 7% sodium dodecyl sulfate (SDS), 10 × Denhardt's, 5% dextran sulfate and 1% salmon sperm DNA; (2) hybridization for 20 hours at 42°C followed by 2 washings of 20 minutes at 20°C in 2 × SSC + 2% SDS, 1 washing of 20 minutes at 20°C in 0.1 × SSC + 0.1% SDS. The final washing is performed in 0.1 × SSC + 0.1% SDS for 30 minutes at 60°C.

The subject of the present invention is also a representative fragment of the polynucleotide as defined above, characterized in that it is capable of being obtained either by the use of restriction enzymes whose recognition and cleavage sites are present in said polynucleotide as defined above, or by amplification with the aid of oligonucleotide primers specific for said polynucleotide as defined above, or by transcription *in vitro*, or by chemical synthesis.

According to an advantageous embodiment of said fragment, it is selected from the group consisting of: the cDNA corresponding to at least one open reading frame (ORF) chosen from: ORF1a, ORF1b, ORF-S, ORF-E,

ORF-M, ORF-N, ORF3, ORF4, ORF7 to ORF11, ORF13 and ORF14 and the cDNA corresponding to the noncoding 5' or 3' ends of said polynucleotide.

According to an advantageous feature of this embodiment, said fragment has a sequence selected from the group consisting of:

- the sequences SEQ ID NO: 2 and 4 representing the cDNA corresponding to the ORF-S which encodes the S protein,
- the sequences SEQ ID NO: 13 and 15 representing the cDNA corresponding to the ORF-E which encodes the E protein,
- the sequences SEQ ID NO: 16 and 18 representing the cDNA corresponding to the ORF-M which encodes the M protein,
- the sequences SEQ ID NO: 36 and 38 representing the cDNA corresponding to the ORF-N which encodes the N protein,
- the sequences representing the cDNA corresponding respectively: to ORF1a and ORF1b (ORF1ab, SEQ ID NO: 31), to ORF3 and ORF4 (SEQ ID NO: 7, 8), to ORF7 to 11 (SEQ ID NO: 19, 20) to ORF13 (SEQ ID NO: 32) and to ORF14 (SEQ ID NO: 34), and
- the sequences representing the cDNAs corresponding respectively to the noncoding 5' (SEQ ID NO: 39 and 72) and 3' (SEQ ID NO: 40, 73) ends of said polynucleotide.

The subject of the present invention is also a cDNA fragment encoding the S protein, as defined above, characterized in that it has a sequence selected from the group consisting of the sequences SEQ ID NO: 5 and 6 (Sa and Sb fragments).

The subject of the present invention is also a cDNA fragment corresponding to ORF1a and ORF1b as defined above, characterized in that it has a sequence selected from the group consisting of the sequences SEQ ID NO: 41 to 54 (L0 to L12 fragments).

The subject of the present invention is also a polynucleotide fragment as defined above, characterized in that it has at least 15 consecutive bases or base pairs of the sequence of the genome of said strain including at least one of those situated in position 7979, 16622, 19064, 23220, 24872, 25298 and 26857. Preferably this is a fragment of 20 to 2500 bases or base pairs, preferably from 20 to 400.

According to an advantageous embodiment of said fragment, it includes at least one pair of bases or base pairs corresponding to the following positions: 7919 and 23220, 7919 and 25298, 16622 and 23220, 19064 and 23220, 16622 and 25298, 19064 and 25298, 23220 and 24872, 23220 and 26857, 24872 and 25298, 25298 and 26857.

The subject of the present invention is also primers of at least 18 bases capable of amplifying a fragment of the genome of a SARS-associated coronavirus or of the DNA equivalent thereof.

According to an embodiment of said primers, they are selected from the group consisting of:

- the pair of primers No. 1 corresponding respectively to positions 28507 to 28522 (sense primer, SEQ ID NO: 60) and 28774 to 28759 (antisense primer, SEQ ID NO: 61) of the sequence of the polynucleotide as defined above,
- the pair of primers No. 2 corresponding respectively to positions 28375 to 28390 (sense primer, SEQ ID NO: 62) and 28702 to 28687 (antisense primer, SEQ ID NO: 63) of the sequence of the polynucleotide as defined above, and
- the pair of primers consisting of the primers SEQ ID Nos: 55 and 56.

The subject of the present invention is also a probe capable of detecting the presence of the genome of a SARS-associated coronavirus or of a fragment thereof, characterized in that it is selected from the group consisting of: the fragments as defined above and the fragments corresponding to the following positions of the polynucleotide sequence as defined above: 28561 to 28586, 28588 to 28608, 28541 to 28563 and 28565 to 28589 (SEQ ID NO: 64 to 67).

The probes and primers according to the invention may be labeled directly or indirectly with a radioactive or nonradioactive compound by methods well known to persons skilled in the art so as to obtain a detectable and/or quantifiable signal. Among the radioactive isotopes used, there may be mentioned ^{32}P, ^{33}P, ^{35}S, ^{3}H or ^{125}I. The nonradioactive entities are selected from ligands such as biotin, avidin, streptavidin, digoxygenin, haptens, dyes, luminescent agents such as radioluminescent, chemoluminescent, bioluminescent, fluorescent and phosphorescent agents.

The invention encompasses the labeled probes and primers derived from the preceding sequences.

Such probes and primers are useful for the diagnosis of infection by a SARS-associated coronavirus.

The subject of the present invention is also a method for the detection of a SARS-associated coronavirus, from a biological sample, which method is characterized in that it comprises at least:

(a) the extraction of nucleic acids present in said biological sample,

(b) the amplification of a fragment of ORF-N by RT-PCR with the aid of a pair of primers as defined above, and

(c) the detection, by any appropriate means, of the amplification products obtained in (b).

The amplification products (amplicons) in (b) are 268 bp for the pair of primers No. 1 and 328 bp for the pair of primers No. 2.

According to an advantageous embodiment of said method, the step (b) of detection is carried out with the aid of at least one probe corresponding to positions 28561 to 28586, 28588 to 28608, 28541 to 28563 and 28565 to 28589 of the sequence of the polynucleotide as defined above.

Preferably, the SARS-associated coronavirus genome is detected and optionally quantified by PCR in real time with the aid of the pair of primers No. 2 and probes corresponding to positions 28541 to 28563 and 28565 to 28589 labeled with different compounds, in particular different fluorescent agents.

The real time RT-PCR which uses this pair of primers and this probe is very sensitive since it makes it possible to detect 10^2 copies of RNA and up to 10 copies of RNA; it is in addition reliable and reproducible.

The invention encompasses the single-stranded, double-stranded and triple-stranded polydeoxyribonucleotides and polyribonucleotides corresponding to the sequence of the genome of the isolated strain of coronavirus and its fragments as defined above, and to their sense or antisense complementary sequences, in particular the RNAs and cDNAs corresponding to the sequence of the genome and of its fragments as defined above.

The present invention also encompasses the amplification fragments obtained with the aid of primers specific for the genome of the purified or isolated strain as defined above, in particular with the aid of primers or pairs of primers as defined above, the restriction fragments formed by or

comprising the sequence of fragments as defined above, the fragments obtained by transcription *in vitro* from a vector containing the sequence SEQ ID NO: 1 or a fragment as defined above, and fragments obtained by chemical synthesis. Examples of restriction fragments are deduced from the restriction map of the sequence SEQ ID NO: 1 illustrated by Figure 13. In accordance with the invention, said fragments are either in the form of isolated fragments, or in the form of mixtures of fragments. The invention also encompasses fragments modified, in relation to the preceding ones, by removal or addition of nucleotides in a proportion of about 15%, relative to the length of the above fragments and/or modified in terms of the nature of the nucleotides, as long as the modified nucleotide fragments retain a capacity for hybridization with the genomic or antigenomic RNA sequences of the isolate as defined above.

The nucleic acid molecules according to the invention are obtained by conventional methods, known per se, following standard protocols such as those described in *Current Protocols in Molecular Biology (Frederick M. AUSUBEL, 2000, Wiley and son Inc., Library of Congress, USA)*. For example, they may be obtained by amplification of a nucleic sequence by PCR or RT-PCR or alternatively by total or partial chemical synthesis.

The subject of the present invention is also a DNA or RNA chip or filter, characterized in that it comprises at least one polynucleotide or one of its fragments as defined above.

The DNA or RNA chips or filters according to the invention are prepared by conventional methods, known per se, such as for example chemical or electrochemical grafting of oligonucleotides on a glass or nylon support.

The subject of the present invention is also a recombinant cloning and/or expression vector, in particular a plasmid, a virus, a viral vector or a phage comprising a nucleic acid fragment as defined above. Preferably, said recombinant vector is an expression vector in which said nucleic acid fragment is placed under the control of appropriate elements for regulating transcription and translation. In addition, said vector may comprise sequences (tags) fused in phase with the 5' and/or 3' end of said insert, which are useful for the immobilization and/or detection and/or purification of the protein expressed from said vector.

These vectors are constructed and introduced into host cells by conventional recombinant DNA and genetic engineering methods which are known per se. Numerous vectors into which a nucleic acid molecule of interest may be inserted in order to introduce it and to maintain it in a host cell are known per se; the choice of an appropriate vector depends on the use envisaged for this vector (for example replication of the sequence of interest, expression of this sequence, maintenance of the sequence in extrachromosomal form or alternatively integration into the chromosomal material of the host), and on the nature of the host cell.

In accordance with the invention, said plasmid is selected in particular from the following plasmids:

- the plasmid, called SARS-S, contained in the bacterial strain deposited under the No. I-3059, on June 20, 2003, at the Collection Nationale de Cultures de Microorganismes, 25 rue du Docteur Roux, 75724 Paris Cedex 15; it contains the cDNA sequence encoding the S protein of the SARS-CoV strain derived from the sample recorded under the No. 031589, said sequence corresponding to the nucleotides at positions 21406 to 25348 (SEQ ID NO: 4), with reference to the Genbank sequence AY274119.3,

- the plasmid, called SARS-S1, contained in the bacterial strain deposited under the No. I-3020, on May 12, 2003, at the Collection Nationale de Cultures de Microorganismes, 25 rue du Docteur Roux, 75724 Paris Cedex 15; it contains a 5' fragment of the cDNA sequence encoding the S protein of the SARS-CoV strain derived from the sample recorded under the No. 031589, as defined above, said fragment corresponding to the nucleotides at positions 21406 to 23454 (SEQ ID NO: 5), with reference to the Genbank sequence AY274119.3 Tor2,
- the plasmid, called SARS-S2, contained in the bacterial strain deposited under the No. I-3019, on May 12, 2003, at the Collection Nationale de Cultures de Microorganismes, 25 rue du Docteur Roux, 75724 Paris Cedex 15; it contains a 3' fragment of the cDNA sequence encoding the S protein of the SARS-CoV strain derived from the sample recorded under the number No. 031589, as defined above, said fragment corresponding to the nucleotides at positions 23322 to 25348 (SEQ ID NO: 6), with reference to the Genbank sequence accession No. AY274119.3,
- the plasmid, called SARS-SE, contained in the bacterial strain deposited under the No. I-3126, on November 13, 2003, at the Collection Nationale de Cultures de Microorganismes, 25 rue du Docteur Roux, 75724 Paris Cedex 15; it contains the cDNA corresponding to the region situated between ORF-S and ORF-E and overlapping ORF-E of the SARS-CoV strain derived from the sample recorded under the No. 031589, as defined above, said region corresponding to the nucleotides at positions 25110 to 26244 (SEQ ID NO: 8), with reference to the Genbank sequence accession No. AY274119.3,
- the plasmid, called SARS-E, contained in the bacterial strain deposited under the No. I-3046, on May 28, 2003, at the Collection Nationale de Cultures de Microorganismes, 25 rue du Docteur Roux, 75724 Paris Cedex 15; it contains the cDNA sequence encoding the E protein of the SARS-CoV strain derived from the sample

recorded under the No. 031589, as defined above, said sequence corresponding to the nucleotides at positions 26082 to 26413 (SEQ ID NO: 15), with reference to the Genbank sequence accession No. AY274119.3,

- the plasmid, called SARS-M, contained in the bacterial strain deposited under the No. I-3047, on May 28, 2003, at the Collection Nationale de Cultures de Microorganismes, 25 rue du Docteur Roux, 75724 Paris Cedex 15; it contains the cDNA sequence encoding the M protein of the SARS-CoV strain derived from the sample recorded under the No. 031589, as defined above; said sequence corresponding to the nucleotides at positions 26330 to 27098 (SEQ ID NO: 18), with reference to the Genbank sequence accession No. AY274119.3,

- the plasmid, called SARS-MN, contained in the bacterial sequence deposited under the No. I-3125, on November 13, 2003, at the Collection Nationale de Cultures de Microorganismes, 25 rue du Docteur Roux, 75724 Paris Cedex 15; it contains the cDNA sequence corresponding to the region situated between ORF-M and ORF-N of the SARS-CoV strain derived from the sample recorded under the No. 031589 and collected in Hanoi, as defined above, said sequence corresponding to the nucleotides at positions 26977 to 28218 (SEQ ID NO: 20), with reference to the Genbank accession No. AY274119.3,

- the plasmid, called SARS-N, contained in the bacterial strain deposited under the No. I-3048, on June 5, 2003, at the Collection Nationale de Cultures de Microorganismes, 25 rue du Docteur Roux, 75724 Paris Cedex 15; it contains the cDNA encoding the N protein of the SARS-CoV strain derived from the sample recorded under the No. 031589, as defined above, said sequence corresponding to the nucleotides at positions 28054 to 29430 (SEQ ID NO: 38), with reference to the Genbank sequence accession No. AY274119.3; thus, this plasmid comprises an insert of sequence SEQ ID NO: 38 and is contained in a bacterial strain which was deposited under the No. I-3048, on June 5, 2003, at the

Collection Nationale de Cultures de Microorganismes, 25 rue du Docteur Roux, 75724 Paris Cedex 15,

- the plasmid, called SARS-5'NC, contained in the bacterial strain deposited under the No. I-3124, on November 7, 2003, at the Collection Nationale de Cultures de Microorganismes, 25 rue du Docteur Roux, 75724 Paris Cedex 15; it contains the cDNA corresponding to the noncoding 5' end of the genome of the SARS-CoV strain derived from the sample recorded under the No. 031589, as defined above, said sequence corresponding to the nucleotides at positions 1 to 204 (SEQ ID NO: 39), with reference to the Genbank sequence accession No. AY274119.3,

- the plasmid called SARS-3'NC, contained in the bacterial strain deposited under the No. I-3123 on November 7, 2003, at the Collection Nationale de Cultures de Microorganismes, 25 rue du Docteur Roux, 75724 Paris Cedex 15; it contains the cDNA sequence corresponding to the noncoding 3' end of the genome of the SARS-CoV strain derived from the sample recorded under the No. 031589, as defined above, said sequence corresponding to that situated between the nucleotide and position 28933 to 29727 (SEQ ID NO: 40), with reference to the Genbank sequence accession No. AY274119.3, ends with a series of nucleotides a.,

- the expression plasmid, called pIV2.3N, containing a cDNA fragment encoding a C-terminal fusion of the N protein (SEQ ID NO: 37) with a polyhistidine tag,

- the expression plasmid, called pIV2.3S_C, containing a cDNA fragment encoding a C-terminal fusion of the fragment corresponding to positions 475 to 1193 of the amino acid sequence of the S protein (SEQ ID NO: 3) with a polyhistidine tag,

- the expression plasmid, pIV2.3S_L, containing a cDNA fragment encoding a C-terminal fusion of the fragment corresponding to positions 14 to 1193 of the amino acid sequence of the S protein (SEQ ID NO: 3) with a polyhistidine tag,

- the expression plasmid, called pIV2.4N, containing a cDNA fragment encoding a N-terminal fusion of the N protein (SEQ ID NO: 3) with a polyhistidine tag,
- the expression plasmid, called pIV2.4S_C or pIV2.4S_1, containing an insert encoding a N-terminal fusion of the fragment corresponding to positions 475 to 1193 of the amino acid sequence of the S protein (SEQ ID NO: 3) with a polyhistidine tag, and
- the expression plasmid, called pIV2.4S_L, containing a cDNA fragment encoding an N-terminal fusion of the fragment corresponding to positions 14 to 1193 of the amino acid sequence of the S protein (SEQ ID NO: 3) with a polyhistidine tag.

According to an advantageous feature of the expression plasmid as defined above, it is contained in a bacterial strain which was deposited under the No. I-3117, on October 23, 2003, at the Collection Nationale de Cultures de Microorganismes, 25 rue du Docteur Roux, 75724 Paris Cedex 15.

According to another advantageous feature of the expression plasmid as defined above, it is contained in a bacterial strain which was deposited under the No. I-3118, on October 23, 2003, at the Collection Nationale de Cultures de Microorganismes, 25 rue du Docteur Roux, 75724 Paris Cedex 15.

According to another feature of the expression plasmid as defined above, it is contained in a bacterial strain which was deposited at the CNCM, 25 rue du Docteur Roux, 75724 Paris Cedex 15 under the following numbers:

a) strain No. I-3118, deposited on October 23, 2003,
b) strain No. I-3019, deposited on May 12, 2003,
c) strain No. I-3020, deposited on May 12, 2003,
d) strain No. I-3059, deposited on June 20, 2003,
e) strain No. I-3323, deposited on November 22, 2004,
f) strain No. I-3324, deposited on November 22, 2004,
g) strain No. I-3326, deposited on December 1, 2004,

h) strain No. I-3327, deposited on December 1, 2004,
i) strain No. I-3332, deposited on December 1, 2004,
j) strain No. I-3333, deposited on December 1, 2004,
k) strain No. I-3334, deposited on December 1, 2004,
l) strain No. I-3335, deposited on December 1, 2004,
m) strain No. I-3336, deposited on December 1, 2004,
n) strain No. I-3337, deposited on December 1, 2004,
o) strain No. I-3338, deposited on December 2, 2004,
p) strain No. I-3339, deposited on December 2, 2004,
q) strain No. I-3340, deposited on December 2, 2004,
r) strain No. I-3341, deposited on December 2, 2004.

The subject of the present invention is also a nucleic acid insert of viral origin, characterized in that it is contained in any of the strains as defined above in a)-r).

The subject of the present invention is also a nucleic acid containing a synthetic gene allowing optimized expression of the S protein in eukaryotic cells, characterized in that it possesses the sequence SEQ ID NO: 140.

The subject of the present invention is also an expression vector containing a nucleic acid containing a synthetic gene allowing optimized expression of the S protein, which vector is contained in the bacterial strain deposited at the CNCM, on December 1, 2004, under the No. I-3333.

According to one embodiment of said expression vector, it is a viral vector, in the form of a viral particle or in the form of a recombinant genome.

According to an advantageous feature of this embodiment, this is a recombinant viral particle or a recombinant viral genome capable of being obtained by transfection of a plasmid according to paragraphs g), h) and k) to r) as defined above, in an appropriate

cellular system, that is to say, for example, cells transfected with one or more other plasmids intended to transcomplement certain functions of the virus that are deleted in the vector and that are necessary for the formation of the viral particles.

The expression "S protein family" is understood here to mean the complete S protein, its ectodomain and fragments of this ectodomain which are preferably produced in a eukaryotic system.

The subject of the present invention is also a lentiviral vector encoding a polypeptide of the S protein family, as defined above.

The subject of the present invention is also a recombinant measles virus encoding a polypeptide of the S protein family, as defined above.

The subject of the present invention is also a recombinant vaccinia virus encoding a polypeptide of the S protein family, as defined above.

The subject of the present invention is also the use of a vector according to paragraphs e) to r) as defined above, or of a vector containing a synthetic gene for the S protein, as defined above, for the production, in a eukaryotic system, of the SARS-associated coronavirus S protein or of a fragment of this protein.

The subject of the present invention is also a method for producing the S protein in a eukaryotic system, comprising a step of transfecting eukaryotic cells in culture with a vector chosen from the vectors contained in the bacterial strains mentioned in paragraphs e) to r) above or a vector containing a synthetic gene allowing optimized expression of the S protein.

The subject of the present invention is also a cDNA library characterized in that it comprises fragments as defined above, in particular amplification fragments or restriction fragments, cloned into a recombinant vector, in particular an expression vector (expression library).

The subject of the present invention is also cells, in particular prokaryotic cells, modified by a recombinant vector as defined above.

The subject of the present invention is also a genetically modified eukaryotic cell expressing a protein or a polypeptide as defined above. Quite obviously, the terms "genetically modified eukaryotic cell" do not denote a cell modified with a wild-type virus.

According to an advantageous embodiment of said cell, it is capable of being obtained by transfection with any of the vectors mentioned in paragraphs K) to N) above.

According to an advantageous feature of this embodiment, this is the cell FRhK4-Ssol-30, deposited at the CNCM on November 22, 2004, under the No. I-3325.

The recombinant vectors as defined above and the cells transformed with said expression vectors are advantageously used for the production of the corresponding proteins and peptides. The expression libraries derived from said vectors, and the cells transformed with said expression libraries are advantageously used to identify the immunogenic epitopes (B and T epitopes) of the SARS-associated coronavirus proteins.

The subject of the present invention is also the purified or isolated proteins and peptides,

characterized in that they are encoded by the polynucleotide or one of its fragments as defined above.

According to an advantageous embodiment of the invention, said protein is selected from the group consisting of:

- the S protein having the sequence SEQ ID NO: 3 or its ectodomain
- the E protein having the sequence SEQ ID NO: 14
- the M protein having the sequence SEQ ID NO: 17
- the N protein having the sequence SEQ ID NO: 37
- the proteins encoded by the ORFs: ORF1a, ORF1b, ORF3, ORF4 and ORF7 to ORF11, ORF13 and ORF14 and having the respective sequence, SEQ ID NO: 74, 75, 10, 12, 22, 24, 26, 28, 30, 33 and 35.

The terms "ectodomain of the S protein" and "soluble form of the S protein" will be used interchangeably below.

According to an advantageous embodiment of the invention, said polypeptide consists of the amino acids corresponding to positions 1 to 1193 of the amino acid sequence of the S protein.

According to another advantageous embodiment of the invention, said peptide is selected from the group consisting of:

a) the peptides corresponding to positions 14 to 1193 and 475 to 1193 of the amino acid sequence of the S protein,

b) the peptides corresponding to positions 2 to 14 (SEQ ID NO: 69) and 100 to 221 of the amino acid sequence of the M protein; these peptides correspond respectively to the ectodomain and to the endodomain of the M protein, and

c) the peptides corresponding to positions 1 to 12 (SEQ ID NO: 70) and 53 to 76 (SEQ ID NO: 71) of the

amino acid sequence of the E protein; these peptides correspond respectively to the ectodomain and to the C-terminal end of the E protein, and

d) the peptides of 5 to 50 consecutive amino acids, preferably of 10 to 30 amino acids, inclusive or partially or completely overlapping the sequence of the peptides as defined in a), b) or c).

The subject of the present invention is also a peptide, characterized in that it has a sequence of 7 to 50 amino acids including an amino acid residue selected from the group consisting of:

- the alanine situated at position 2552 of the amino acid sequence of the protein encoded by ORF1a,
- the serine situated at position 577 of the amino acid sequence of the S protein of the SARS-CoV strain as defined above,
- the glycine at position 11 of the amino acid sequence of the protein encoded by ORF3 of the SARS-CoV strain as defined above,
- the serine at position 154 of the amino acid sequence of the M protein of the SARS-CoV strain as defined above.

The subject of the present invention is also an antibody or a polyclonal or monoclonal antibody fragment which can be obtained by immunization of an animal with a recombinant vector as defined above, a cDNA library as defined above or alternatively a protein or a peptide as defined above, characterized in that it binds to at least one of the proteins encoded by SARS-CoV as defined above.

The invention encompasses the polyclonal antibodies, the monoclonal antibodies, the chimeric antibodies such as the humanized antibodies, and fragments thereof (Fab, Fv, scFv).

A subject of the present invention is also a hybridoma producing a monoclonal antibody against the N protein, characterized in that it is chosen from the following hybridomas:

- the hybridoma producing the monoclonal antibody 87, deposited at the CNCM on December 1, 2004 under the number I-3328,
- the hybridoma producing the monoclonal antibody 86, deposited at the CNCM on December 1, 2004 under the number I-3329,
- the hybridoma producing the monoclonal antibody 57, deposited at the CNCM on December 1, 2004 under the number I-3330, and
- the hybridoma producing the monoclonal antibody 156, deposited at the CNCM on December 1, 2004 under the number I-3331.

The subject of the present invention is also a polyclonal or monoclonal antibody or antibody fragment directed against the N protein, characterized in that it is produced by a hybridoma as defined above.

For the purposes of the present invention, the expression chimeric antibody is understood to mean, in relation to an antibody of a particular animal species or of a particular class of antibody, an antibody comprising all or part of a heavy chain and/or of a light chain of an antibody of another animal species or of another class of antibody.

For the purposes of the present invention, the expression humanized antibody is understood to mean a human immunoglobulin in which the residues of the CDRs (*Complementary Determining Regions*) which form the antigen-binding site are replaced by those of a nonhuman monoclonal antibody possessing the desired specificity, affinity or activity. Compared with the nonhuman antibodies, the humanized antibodies are less immunogenic and possess a prolonged half-life in humans

because they possess only a small proportion of nonhuman sequences given that practically all the residues of the FR (Framework) regions and of the constant (Fc) region of these antibodies are those of a consensus sequence of human immunoglobulins.

A subject of the present invention is also a protein chip or filter, characterized in that it comprises a protein, a peptide or alternatively an antibody as defined above.

The protein chips according to the invention are prepared by conventional methods known per se. Among the appropriate supports on which proteins may be immobilized, there may be mentioned those made of plastic or glass, in particular in the form of microplates.

The subject of the present invention is also reagents derived from the isolated strain of SARS-associated coronavirus, derived from the sample recorded under the No. 031589, which are useful for the study and diagnosis of the infection caused by a SARS-associated coronavirus, said reagents are selected from the group consisting of:

(a) a pair of primers, a probe or a DNA chip as defined above,

(b) a recombinant vector or a modified cell as defined above,

(c) an isolated coronavirus strain or a polynucleotide as defined above,

(d) a protein or a peptide as defined above,

(e) an antibody or an antibody fragment as defined above, and

(f) a protein chip as defined above.

These various reagents are prepared and used according to conventional molecular biology and immunology techniques following standard protocols such as those

described in *Current Protocols in Molecular Biology* (Frederick M. AUSUBEL, 2000, Wiley and Son Inc., Library of Congress, USA), in *Current Protocols in Immunology* (John E. Cologan, 2000, Wiley and Son Inc., Library of Congress, USA) and in *Antibodies: A Laboratory Manual* (E. Howell and D. Lane, Cold Spring Harbor Laboratory, 1988).

The nucleic acid fragments according to the invention are prepared and used according to conventional techniques as defined above. The peptides and proteins according to the invention are prepared by recombinant DNA techniques, known to persons skilled in the art, in particular with the aid of the recombinant vectors as defined above. Alternatively, the peptides according to the invention may be prepared by conventional techniques of solid or liquid phase synthesis, known to persons skilled in the art.

The polyclonal antibodies are prepared by immunizing an appropriate animal with a protein or a peptide as defined above, optionally coupled to KLH or to albumin and/or combined with an appropriate adjuvant such as (complete or incomplete) Freund's adjuvant or aluminum hydroxide; after obtaining a satisfactory antibody titer, the antibodies are harvested by collecting serum from the immunized animals and enriched with IgG by precipitation, according to conventional techniques, and then the IgGs specific for the SARS-CoV proteins are optionally purified by affinity chromatography on an appropriate column to which said peptide or said protein is attached, as defined above, so as to obtain a monospecific IgG preparation.

The monoclonal antibodies are produced from hybridomas obtained by fusion of B lymphocytes from an animal immunized with a protein or a peptide as defined above with myelomas, according to the Köhler and Milstein technique (Nature, 1975, 256, 495-497); the hybridomas

are cultured *in vitro*, in particular in fermenters or produced *in vivo*, in the form of ascites; alternatively, said monoclonal antibodies are produced by genetic engineering as described in American patent US 4,816,567.

The humanized antibodies are produced by general methods such as those described in International application WO 98/45332.

The antibody fragments are produced from the cloned V_H and V_L regions, from the mRNAs of hybridomas or splenic lymphocytes of an immunized mouse; for example, the Fv, scFv or Fab fragments are expressed at the surface of filamentous phages according to the Winter and Milstein technique (Nature, 1991, 349, 293-299); after several selection steps, the antibody fragments specific for the antigen are isolated and expressed in an appropriate expression system, by conventional techniques for cloning and expression of recombinant DNA.

The antibodies or fragments thereof as defined above are purified by conventional techniques known to persons skilled in the art, such as affinity chromatography.

The subject of the present invention is additionally the use of a product selected from the group consisting of: a pair of primers, a probe, a DNA chip, a recombinant vector, a modified cell, an isolated coronavirus strain, a polynucleotide, a protein or a peptide, an antibody or an antibody fragment and a protein chip as defined above, for the preparation of a reagent for the detection and optionally genotyping/serotyping of a SARS-associated coronavirus.

The proteins and peptides according to the invention, which are capable of being recognized and/or of

inducing the production of antibodies specific for the SARS-associated coronavirus, are useful for the diagnosis of infection with such a coronavirus; the infection is detected, by an appropriate technique - in particular EIA, ELISA, RIA, immunofluorescence -, in a biological sample collected from an individual capable of being infected.

According to an advantageous feature of said use, said proteins are selected from the group consisting of the S, E, M and/or N proteins and the peptides as defined above.

The S, E, M and/or N proteins and the peptides derived from these proteins as defined above, for example the N protein, are used for the indirect diagnosis of a SARS-associated coronavirus infection (serological diagnosis; detection of an antibody specific for SARS-CoV), in particular by an immunoenzymatic method (ELISA).

The antibodies and antibody fragments according to the invention, in particular those directed against the S, E, M and/or N proteins and the derived peptides as defined above, are useful for the direct diagnosis of a SARS-associated coronavirus infection; the detection of the protein(s) of SARS-CoV is carried out by an appropriate technique, in particular EIA, ELISA, RIA, immunofluorescence, in a biological sample collected from an individual capable of being infected.

The subject of the present invention is also a method for the detection of a SARS-associated coronavirus, from a biological sample, which method is characterized in that it comprises at least:

(a) bringing said biological sample into contact with at least one antibody or one antibody fragment, one protein, one peptide or alternatively one protein or peptide chip or filter as defined above, and

(b) visualizing by any appropriate means antigen-antibody complexes formed in (a), for example by EIA, ELISA, RIA, or by immunofluorescence.

According to one advantageous embodiment of said process, step (a) comprises:

(a_1) bringing said biological sample into contact with at least a first antibody or an antibody fragment which is attached to an appropriate support, in particular a microplate,

(a_2) washing the solid phase, and

(a_3) adding at least a second antibody or an antibody fragment, different from the first, said antibody or antibody fragment being optionally appropriately labeled.

This method, which makes it possible to capture the viral particles present in the biological sample, is also called immunocapture method.

For example:

- step (a_1) is carried out with at least a first monoclonal or polyclonal antibody or a fragment thereof, directed against the S, M and/or E protein, and/or a peptide corresponding to the ectodomain of one of these proteins (M2-14 or E1-12 peptides)

- step (a_3) is carried out with at least one antibody or an antibody fragment directed against another epitope of the same protein or preferably against another protein, preferably against an inner protein such as the N nucleoprotein or the endodomain of the E or M protein, more preferably still these are antibodies or antibody fragments directed against the N protein which is very abundant in the viral particle; when an antibody or an antibody fragment directed against an inner protein (N) or against the endodomain of the E or M proteins is used, said antibody is incubated in the presence of detergent, such as Tween 20 for example, at concentrations of the order of 0.1%.

- step (b) for visualizing the antigen-antibody complexes formed is carried out, either directly with the aid of a second antibody labeled for example with biotin or an appropriate enzyme such as peroxidase or alkaline phosphatase, or indirectly with the aid of an anti-immunoglobulin serum labeled as above. The complexes thus formed are visualized with the aid of an appropriate substrate.

According to a preferred embodiment of this aspect of the invention, the biological sample is mixed with the visualizing monoclonal antibody prior to its being brought into contact with the capture monoclonal antibodies. Where appropriate, the serum-visualizing antibody mixture is incubated for at least 10 minutes at room temperature before being applied to the plate.

The subject of the present invention is also an immunocapture test intended to detect an infection by the SARS-associated coronavirus by detecting the native nucleoprotein (N protein), in particular characterized in that the antibody used for the capture of the native viral nucleoprotein is a monoclonal antibody specific for the central region and/or for a conformational epitope.

According to one embodiment of said test, the antibody used for the capture of the N protein is the monoclonal antibody mAb87, produced by the hybridoma deposited at the CNCM on December 1, 2004 under the number I-3328.

According to another embodiment of said immunocapture test, the antibody used for the capture of the N protein is the monoclonal antibody mAb86, produced by the hybridoma deposited at the CNCM on December 1, 2004 under the number I-3329.

According to another embodiment of said immunocapture test, the monoclonal antibodies mAb86 and mAb87 are used for the capture of the N protein.

In the immunocapture tests according to the invention, it is possible to use, for visualizing the N protein, the monoclonal antibody mAb57, produced by the hybridoma deposited at the CNCM on December 1, 2004 under the number I-3330, said antibody being conjugated with a visualizing molecule or particle.

In accordance with said immunocapture test, a combination of the antibodies mAb57 and mAb87, conjugated with a visualizing molecule or particle, is used for the visualization of the N protein.

A visualizing molecule may be a radioactive atom, a dye, a fluorescent molecule, a fluorophore, an enzyme; a visualizing particle may be for example: colloidal gold, a magnetic particle or a latex bead.

The subject of the present invention is also a reagent for detecting a SARS-associated coronavirus, characterized in that it is selected from the group consisting of:

(a) a pair of primers or a probe as defined above,

(b) a recombinant vector as defined above or a modified cell as defined above,

(c) an isolated coronavirus strain as defined above or a polynucleotide as defined above,

(d) an antibody or an antibody fragment as defined above,

(e) a combination of antibodies comprising the monoclonal antibodies mAb86 and/or mAb87, and the monoclonal antibody mAb57, as defined above,

(f) a chip or a filter as defined above.

The subject of the present invention is also a method for the detection of a SARS-associated coronavirus

infection, from a biological sample, by indirect IgG ELISA using the N protein, which method is characterized in that the plates are sensitized with an N protein solution at a concentration of between 0.5 and 4 µg/ml, preferably to 2 µg/ml, in a 10 mM PBS buffer pH 7.2, phenol red at 0.25 ml/l.

The subject of the present invention is additionally a method for the detection of a SARS-associated coronavirus infection, from a biological sample, by double epitope ELSA, characterized in that the serum to be tested is mixed with the visualizing antigen, said mixture then being brought into contact with the antigen attached to a solid support.

According to one variant of the tests for detecting SARS-associated coronaviruses, these tests combine an ELSA using the N protein, and another ELSA using the S protein, as described below.

The subject of the present invention is also an immune complex formed of a polyclonal or monoclonal antibody or antibody fragment as defined above, and of a SARS-associated coronavirus protein or peptide.

The subject of the present invention is additionally a SARS-associated coronavirus detection kit, characterized in that it comprises at least one reagent selected from the group consisting of: a pair of primers, a probe, a DNA or RNA chip, a recombinant vector, a modified cell, an isolated coronavirus strain, a polynucleotide, a protein or a peptide, an antibody, and a protein chip as defined above.

The subject of the present invention is additionally an immunogenic composition, characterized in that it comprises at least one product selected from the group consisting of:

a) a protein or a peptide as defined above,

b) a polynucleotide of the DNA or RNA type or one of its representative fragments as defined above, having a sequence chosen from:

(i) the sequence SEQ ID NO: 1 or its RNA equivalent

(ii) the sequence hybridizing under high stringency conditions with the sequence SEQ ID NO: 1,

(iii) the sequence complementary to the sequence SEQ ID NO: 1 or to the sequence hybridizing under high stringency conditions with the sequence SEQ ID NO: 1,

(iv) the nucleotide sequence of a representative fragment of the polynucleotide as defined in (i), (ii) or (iii),

(v) the sequence as defined in (i), (ii), (iii) or (iv), modified, and

c) a recombinant expression vector comprising a polynucleotide as defined in b), and

d) a cDNA library as defined above,

said immunogenic composition being capable of inducing protective humoral or cellular immunity specific for the SARS-associated coronavirus, in particular the production of an antibody directed against a specific epitope of the SARS-associated coronavirus.

The proteins and peptides as defined above, in particular the S, M, E and/or N proteins and the derived peptides, and the nucleic acid (DNA or RNA) molecules encoding said proteins or said peptides are good candidate vaccines and may be used in immunogenic compositions for the production of a vaccine against the SARS-associated coronavirus.

According to an advantageous embodiment of the compositions according to the invention, they additionally contain at least one pharmaceutically acceptable vehicle and optionally carrier substances and/or adjuvants.

The pharmaceutically acceptable vehicles, the carrier substances and the adjuvants are those conventionally used.

The adjuvants are advantageously chosen from the group consisting of oily emulsions, saponin, mineral substances, bacterial extracts, aluminum hydroxide and squalene.

The carrier substances are advantageously selected from the group consisting of unilamellar liposomes, multilamellar liposomes, micelles of saponin or solid microspheres of a saccharide or auriferous nature.

The compositions according to the invention are administered by the general route, in particular by the intramuscular or subcutaneous route or alternatively by the local, in particular nasal (aerosol) route.

The subject of the present invention is also the use of an isolated or purified protein or peptide having a sequence selected from the group consisting of the sequences SEQ ID NO: 3, 10, 12, 14, 17, 22, 24, 26, 28, 30, 33, 35, 37, 69, 70, 71, 74 and 75 to form an immune complex with an antibody specifically directed against an epitope of the SARS-associated coronavirus.

The subject of the present invention is also an immune complex consisting of an isolated or purified protein or peptide having a sequence selected from the group consisting of the sequences SEQ ID NO: 3, 10, 12, 14, 17, 22, 24, 26, 28, 30, 33, 35, 37, 69, 70, 71, 74 and 75, and of an antibody specifically directed against an epitope of the SARS-associated coronavirus.

The subject of the present invention is also the use of an isolated or purified protein or peptide having a sequence selected from the group consisting of the sequences SEQ ID NO: 3, 10, 12, 14, 17, 22, 24, 26, 28,

30, 33, 35, 37, 69, 70, 71, 74 and 75 to induce the production of an antibody capable of specifically recognizing an epitope of the SARS-associated coronavirus.

The subject of the present invention is also the use of an isolated or purified polynucleotide having a sequence selected from the group consisting of the sequences SEQ ID NO: 1, 2, 4, 7, 8, 13, 15, 16, 18, 19, 20, 31, 36 and 38 to induce the production of an antibody directed against the protein encoded by said polynucleotide and capable of specifically recognizing an epitope of the SARS-associated coronavirus.

The subject of the present invention is also monoclonal antibodies recognizing the native S protein of a SARS-associated coronavirus.

The subject of the present invention is also the use of a protein or a polypeptide of the S protein family, as defined above, or of an antibody recognizing the native S protein, as defined above, to detect an infection by a SARS-associated coronavirus, in a biological sample.

The subject of the present invention is also a method for detecting an infection by a SARS-associated coronavirus, in a biological sample, characterized in that the detection is carried out by ELISA using the recombinant S protein, expressed in a eukaryotic system.

According to an advantageous embodiment of said method, it is a double epitope ELISA method, and the serum to be tested is mixed with the visualizing antigen, said mixture then being brought into contact with the antigen attached to a solid support.

The subject of the present invention is also an immune complex consisting of a monoclonal antibody or antibody

fragment recognizing the native S protein, and of a protein or a peptide of the SARS-associated coronavirus.

The subject of the present invention is also an immune complex consisting of a protein or a polypeptide of the S protein family, as defined above, and of an antibody specifically directed against an epitope of the SARS-associated coronavirus.

The subject of the present invention is additionally a SARS-associated coronavirus detection kit or box, characterized in that it comprises at least one reagent selected from the group consisting of: a protein or polypeptide of the S protein family, as defined above, a nucleic acid encoding a protein or peptide of the S protein family, as defined above, a cell expressing a protein or polypeptide of the S protein family, as defined above, or an antibody recognizing the native S protein of a SARS-associated coronavirus.

The subject of the present invention is an immunogenic and/or vaccine composition, characterized in that it comprises a polypeptide or a recombinant protein of the S protein family, as defined above, obtained in a eukaryotic expression system.

The subject of the present invention is also an immunogenic and/or vaccine composition, characterized in that it comprises a vector or recombinant virus, expressing a protein or a polypeptide of the S protein family, as defined above.

In addition to the preceding features, the invention further comprises other features, which will emerge from the description which follows, which refers to examples of use of the polynucleotide representing the genome of the SARS-CoV strain derived from the sample recorded under the number 031589, and derived cDNA

fragments which are the subject of the present invention, and to Table I presenting the sequence listing:

Table I: Sequence listing

Identification number	Sequence	Position of the cDNA with reference to Genbank AY274119.3	Deposit number at the CNCM of the corresponding plasmid
SEQ ID NO : 1	genome of the strain derived from the sample 031589	-	-
SEQ ID NO : 2	ORF-S*	21406-25348	-
SEQ ID NO : 3	S protein	-	-
SEQ ID NO : 4	ORF-S**	21406-25348	I-3059
SEQ ID NO : 5	Sa fragment	21406-23454	I-3020
SEQ ID NO : 6	Sb fragment	23322-25348	I-3019
SEQ ID NO : 7	ORF-3+ORF-4*	25110-26244	-
SEQ ID NO : 8	ORF-3+ORF-4**	25110-26244	I-3126
SEQ ID NO : 9	ORF3	-	-
SEQ ID NO : 10	ORF-3 protein	-	-
SEQ ID NO : 11	ORF4	-	-
SEQ ID NO : 12	ORF-4 protein	-	-
SEQ ID NO : 13	ORF-E*	26082-26413	-
SEQ ID NO : 14	E protein	-	-
SEQ ID NO : 15	ORF-E**	26082-26413	I-3046
SEQ ID NO : 16	ORF-M*	26330-27098	-
SEQ ID NO : 17	M protein	-	-
SEQ ID NO : 18	ORF-M**	26330-27098	I-3047
SEQ ID NO : 19	ORF7 to 11*	26977-28218	-
SEQ ID NO : 20	ORF7 to 11**	26977-28218	I-3125
SEQ ID NO : 21	ORF7	-	-
SEQ ID NO : 22	ORF7 protein	-	-
SEQ ID NO : 23	ORF8	-	-
SEQ ID NO : 24	ORF8 protein	-	-
SEQ ID NO : 25	ORF9	-	-

SEQ ID NO : 26	ORF9 protein	-	-
SEQ ID NO : 27	ORF10	-	-
SEQ ID NO : 28	ORF10 protein	-	-
SEQ ID NO : 29	ORF11	-	-
SEQ ID NO : 30	ORF11 protein	-	-
SEQ ID NO: 31	OrF1ab	265-21485	-
SEQ ID NO : 32	ORF13	28130-28426	-
SEQ ID NO : 33	ORF13 protein	-	-
SEQ ID NO : 34	ORF14	-	-
SEQ ID NO : 35	ORF14 protein	28583-28795	-
SEQ ID NO : 36	ORF-N*	28054-29430	
SEQ ID NO : 37	N protein	-	-
SEQ ID NO : 38	ORF-N**	28054-29430	I-3048
SEQ ID NO : 39	noncoding 5'**	1-204	I-3124
SEQ ID NO : 40	noncoding 3'**	28933-29727	I-3123
SEQ ID NO : 41	**ORF1ab** Fragment L0	30-500	-
SEQ ID NO : 42	Fragment L1	211-2260	-
SEQ ID NO : 43	Fragment L2	2136-4187	-
SEQ ID NO : 44	Fragment L3	3892-5344	-
SEQ ID NO : 45	Fragment L4b	4932-6043	-
SEQ ID NO : 46	Fragment L4	5305-7318	-
SEQ ID NO : 47	Fragment L5	7275-9176	-
SEQ ID NO : 48	Fragment L6	9032-11086	-
SEQ ID NO : 49	Fragment L7	10298-12982	-
SEQ ID NO : 50	Fragment L8	12815-14854	-
SEQ ID NO : 51	Fragment L9	14745-16646	-
SEQ ID NO : 52	Fragment L10	16514-18590	-
SEQ ID NO : 53	Fragment L11	18500-20602	-
SEQ ID NO : 54	Fragment L12	20319-22224	-
SEQ ID NO : 55	Sense N primer	-	-
SEQ ID NO : 56	Antisense N primer	-	-
SEQ ID NO : 57	Sense S_C primer	-	-
SEQ ID NO : 58	Sense S_L primer	-	-
SEQ ID NO : 59	Antisense S_C and S_L primer	-	-
SEQ ID NO : 60	Sense primer series 1	28507-28522	-
SEQ ID NO : 61	Antisense primer series 1	28774-28759	
SEQ ID NO : 62	Sense primer series 2	28375-28390	-
SEQ ID NO : 63	Antisense primer series 2	28702-28687	-
SEQ ID NO : 64	Probe 1/series 1	28561-28586	-

SEQ ID NO : 65	Probe 2/series 1	28588-28608	-
SEQ ID NO : 66	Probe 1/series 2	28541-28563	-
SEQ ID NO : 67	Probe 2/series 2	28565-28589	-
SEQ ID NO : 68	Anchor primer 14T		
SEQ ID NO : 69	Peptide M2-14	-	-
SEQ ID NO : 70	Peptide E1-12	-	-
SEQ ID NO : 71	Peptide E53-76	-	-
SEQ ID NO : 72	Noncoding 5'*	1-204	-
SEQ ID NO : 73	Noncoding 3'*	28933-29727	-
SEQ ID NO : 74	ORF1a protein	-	-
SEQ ID NO : 75	ORF1b protein	-	-
SEQ ID NO:76-139	Primers		
SEQ ID NO:140	Pseudogene of S		
SEQ ID NO:141-148	Primers		
SEQ ID NO:149	Aa1-13 of S		
SEQ ID NO:150	Polypeptide		
SEQ ID NO:151-158	Primers		

* PCR amplification product (amplicon)

** Insert cloned into the plasmid deposited at the CNCM

and to the appended drawings in which:

- Figure 1 illustrates Western-blot analysis of the expression *in vitro* of the recombinant proteins N, S_C and S_L from the expression vectors pIVEX. Lane 1: pIV2.3N. Lane 2: pIV2.3S_C. Lane 3: pIV2.3S_L. Lane 4: pIV2.4N. Lane 5: pIV2.4S_1 or pIV2.4S_C. Lane 6: pIV2.4S_L. The expression of the GFP protein expressed from the same vector is used as a control.

- Figure 2 illustrates the analysis, by polyacrylamide gel electrophoresis under denaturing conditions (SDS-PAGE) and staining with Coomassie blue, of the expression *in vivo* of the N protein from the expression vectors pIVEX. The *E. coli* BL21(DE3)pDIA17 strain transformed with the recombinant vectors pIVEX is cultured at 30°C in LB medium, in the presence or in the absence of inducer (IPTG 1 mM). Lane 1: pIV2.3N. Lane 2: pIV2.4N.

- Figure 3 illustrates the analysis, by polyacrylamide gel electrophoresis under denaturing conditions

(SDS-PAGE) and staining with Coomassie blue, of the expression *in vivo* of the S_L and S_C polypeptides from the expression vectors pIVEX. The *E. coli* BL21(DE3)pDIA17 strain transformed with the recombinant vectors pIVEX is cultured at 30°C in LB medium, in the presence or in the absence of inducer (IPTG 1 mM). Lane 1: pIV2.3S_C. Lane 2: pIV2.3S_L. Lane 3: pIV2.4S_1. Lane 4: pIV2.4S_L.

- Figure 4 illustrates the antigenic activity of the recombinant N, S_L and S_C proteins produced in the *E. coli* BL21(DE3)pDIA17 strain transformed with the recombinant vectors pIVEX. A: electrophoresis (SDS-PAGE) of the bacterial lysates. B and C: Western-blot with the sera, obtained from the same patient infected with SARS-CoV, collected 8 days (B: serum M12) and 29 days (C: serum M13) respectively after the onset of the SARS symptoms. Lane 1: pIV2.3N. Lane 2: pIV2.4N. Lane 3: pIV2.3S_C. Lane 4: pIV2.4S_1. Lane 5: pIV2.3S_L. Lane 6: pIV2.4S_L.

- Figure 5 illustrates the purification on an Ni-NTA agarose column of the recombinant N protein produced in the *E. coli* BL21(DE3)pDIA17 strain from the vector pIV2.3N. Lane 1: total bacterial extract. Lane 2: soluble extract. Lane 3: insoluble extract. Lane 4: extract deposited on the Ni-NTA column. Lane 5: unbound proteins. Lane 6: fractions of peak 1. Lane 7: fractions of peak 2.

- Figure 6 illustrates the purification of the recombinant S_C protein from the inclusion bodies produced in the *E. coli* BL21(DE3)pDIA17 strain transformed with pIV2.4S_1. A. Treatment with Triton X-100 (2%): Lane 1: total bacterial extract. Lane 2: soluble extract. Lane 3: insoluble extract. Lane 4: supernatant after treatment with Triton X-100 (2%). Lanes 5 and 6: pellet after treatment with Triton X-100

(2%). B: Treatment with 4 M, 5 M, 6 M and 7 M urea of the soluble and insoluble extracts.

- Figure 7 represents the immunoblot produced with the aid of a lysate of cells infected with SARS-CoV and a serum from a patient suffering from atypical pneumopathy.

- Figure 8 represents immunoblots produced with the aid of a lysate of cells infected with SARS-CoV and rabbit immunosera specific for the nucleoprotein N (A) and for the spicule protein S (B). I.S.: immune serum. p.i.: preimmune serum. The anti-N immune serum was used at 1/50 000 and the anti-S immune serum at 1/10 000.

- Figure 9 illustrates the ELISA reactivity of the rabbit monospecific polyclonal sera directed against the N protein or the short fragment of the S protein (S_c), toward the corresponding recombinant proteins used for immunization. A: rabbits P13097, P13081 and P13031 immunized with the purified recombinant N protein. B: rabbits P11135, P13042 and P14001 immunized with a preparation of inclusion bodies corresponding to the short fragment of the S protein (S_c). I.S.: immune serum. p.i.: preimmune serum.

- Figure 10 illustrates the ELISA reactivity of the purified recombinant N protein, toward sera from patients suffering from atypical pneumonia caused by SARS-CoV. Figure 10a: ELISA plates prepared with the N protein at the concentration of 4 µg/ml and 2 µg/ml. Figure 10B: ELISA plate prepared with the N protein at the concentration of 1 µg/ml. The sera designated A, B, D, E, F, G, H correspond to those of Table IV.

- Figure 11 illustrates the amplification by RT-PCR of decreasing quantities of synthetic RNA of the SARS-CoV N gene (10^7 to 1 copy), with the aid of pairs of primers No. 1 (N/+/28507, N/-/28774) (A) and No. 2

(N/+/28375, N/-/28702) (B). T: amplification performed in the absence of RNA. MW: DNA marker.

- Figure 12 illustrates the amplification by RT-PCR in real time of synthetic RNA for the SARS-CoV N gene: decreasing quantities of synthetic RNA as replica (repli.; lanes 16 to 29) and of viral RNA diluted $1/20 \times 10^{-4}$ (lane 32) were amplified by RT-PCR in real time with the aid of the kit "Light Cycler RNA Amplification Kit Hybridization Probes" and pairs of primers and probes of the No. 2 series, under the conditions described in Example 8.

- Figure 13 (Figures 13.1 to 13.7) represents the restriction map of the sequence SEQ ID NO: 1 corresponding to the DNA equivalent of the genome of the SARS-CoV strain derived from the sample recorded under the number 031589.

- Figure 14 shows the result of the SARS serology test by indirect N ELISA (1^{st} series of sera tested).

- Figure 15 shows the result of the SARS serology test by indirect N ELISA (2^{nd} series of sera tested).

- Figure 16 presents the result of the SARS serology test by double epitope N ELISA (1^{st} series of sera tested).

- Figure 17 shows the result of the SARS serology test by double epitope N ELISA (2^{nd} series of sera tested).

- Figure 18 illustrates the test of reactivity of the anti-N monoclonal antibodies by ELISA on the native nucleoprotein N of SARS-CoV. The antibodies were tested in the form of hybridoma culture supernatants by indirect ELISA using an irradiated lysate of VeroE6 cells infected with SARS-CoV as antigen (SARS lysate curves). A negative control for reactivity is performed

for each antibody on a lysate of uninfected VeroE6 cells (negative lysate curves). Several monoclonal antibodies of known specificity were used as negative control antibodies: para1-3 directed against the antigens of the parainfluenza viruses type 1-3 (Bio-Rad) and influenza B directed against the antigens of the influenza virus type B (Bio-Rad).

- Figure 19 illustrates the test of reactivity of the anti-N of SARS-CoV monoclonal antibodies by ELISA on the native antigens of the human coronavirus 229E (HCoV-229E). The antibodies were tested in the form of hybridoma culture supernatants by an indirect ELISA test using a lysate of MRC-5 cells infected with the human coronavirus 229E as antigen (229E lysate curves). A negative control for immunoreactivity was performed for each antibody on a lysate of noninfected MRC-5 cells (negative lysate curves). The monoclonal antibody 5-11H.6 directed against the S protein of the human coronavirus 229E (Sizun et al. 1998, J. Virol. Met. 72: 145-152) is used as positive control antibody. The antibodies para1-3 directed against the antigens of the parainfluenza virus type 1-3 (Bio-Rad) and influenza B directed against the antigens of the influenza virus type B (Bio-Rad) were added to the panel of monoclonal antibodies tested.

- Figure 20 shows a test of reactivity of the anti-N of SARS-CoV monoclonal antibodies by Western blotting on the denatured native nucléoprotein N of SARS-CoV. A lysate of VeroE6 cells infected with SARS-CoV was prepared in the loading buffer according to Laemmli and caused to migrate in a 12% SDS polyacrylamide gel and then the proteins were transferred onto PVDF membrane. The anti-N monoclonal antibodies tested were used for the immunoassay at the concentration of 0.05 µg/ml. The visualization is carried out with anti-mouse IgG(H+L) antibodies coupled to peroxidase (NA93IV, Amersham) and the ECL+ system. Two monoclonal antibodies were used as

negative controls for reactivity: influenza B directed against the antigens of the influenza virus type B (Bio-Rad) and para1-3 directed against the antigens of the parainfluenza virus type 1-3 (Bio-Rad).

- Figure 21 presents the plasmids for expression in mammalian cells of the SARS-CoV S protein. The cDNA for the SARS-CoV S was inserted between the BamH1 and Xho1 sites of the expression plasmid pcDNA3.1(+) (Clontech) in order to obtain the plasmid pcDNA-S and between the Nhe1 and Xho1 sites of the expression plasmid pCI (Promega) in order to obtain the plasmid pCI-S. The WPRE and CTE sequences were inserted between each of the two plasmids pcDNA-S and pCI-S between the Xho1 and Xba1 sites in order to obtain the plasmids pcDNA-S-CTE, pcDNA-S-WPRE, pCI-S-CTE and pCI-S-WPRE, respectively.

SP: signal peptide predicted (aa 1-13) with the software signalP v2.0 (Nielsen et al., 1997, Protein Engineering, 10: 1-6)

TM: transmembrane region predicted (aa 1196-1218) with the software TMHMM v2.0 (Sonnhammer et al., 1998, Proc. of Sixth Int. Conf. on Intelligent Systems for Molecular Biology, pp. 175-182, AAAI Press). It should be noted that the amino acids W1194 and P1195 are possibly part of the transmembrane region with the respective probabilities of 0.13 and 0.42

P-CMV: cytomegalovirus immediate/early promoter.
BGH pA: polyadenylation signal of the bovine growth hormone gene

SV40 late pA: SV40 virus late polyadenylation signal

SD/SA: splice donor and acceptor sites

WPRE: sequences of the "Woodchuck Hepatitis Virus posttranscriptional regulatory element" of the woodchuck hepatitis virus

CTE: sequences of the "constitutive transport element" of the Mason-Pfizer simian retrovirus

- Figure 22 illustrates the expression of the S protein after transfection of VeroE6 cells. Cellular extracts were prepared 48 hours after transfection of VeroE6 cells with the plasmids pcDNA, pcDNA-S, pCI and pCI-S. Cellular extracts were also prepared 18 hours after infection with the recombinant vaccinia virus VV-TF7.3 and transfection with the plasmids pcDNA or pcDNA-S. As a control, extracts of VeroE6 cells were prepared 8 hours after infection with SARS-CoV at a multiplicity of infection of 3. They were separated on an 8% SDS acrylamide gel and analyzed by Western blotting with the aid of an anti-S rabbit polyclonal antibody and an anti-rabbit IgG(H+L) polyclonal antibody coupled to peroxidase (NA934V, Amersham). A molecular mass ladder (kDa) is presented in the figure.

SARS-CoV: extract of VeroE6 cells infected with SARS-CoV

Mock: control extract of noninfected cells

- Figure 23 illustrates the effect of the CTE and WPRE sequences on the expression of the S protein after transfection of VeroE6 and 293T cells. Cellular extracts were prepared 48 hours after transfection of VeroE6 cells (A) or 293T cells (B) with the plasmids pcDNA, pcDNA-S, pcDNA-S-CTE, pcDNA-S-WPRE, pCI-S, pCI-S-CTE and pCI-S-WPRE separated on 8% SDS polyacrylamide gel and analyzed by Western blotting with the aid of an anti-S rabbit polyclonal antibody and an anti-rabbit IgG(H+L) polyclonal antibody coupled to peroxidase (NA934V, Amersham). A molecular mass ladder (kDa) is presented in the figure.

SARS-CoV: extract of VeroE6 cells prepared 8 hours after infection with SARS-CoV at a multiplicity of infection of 3.

Mock: control extract of noninfected VeroE6 cells

- Figure 24 presents defective lentiviral vectors with central DNA flap for the expression of SARS-CoV S. The cDNA for the SARS-CoV S protein was cloned in the form

of a BamH1-Xho1 fragment into the plasmid pTRIPΔU3-CMV containing a defective lentiviral vector TRIP with central DNA flap (Sirven et al., 2001, Mol. Ther., 3: 438-448) in order to obtain the plasmid pTRIP-S. The optimum expression cassettes consisting of the CMV virus immediate/early promoter, a splice signal, cDNA for S and either of the posttranscriptional signals CTE or WPRE were substituted for the cassette EF1α-EGFP of the defective lentiviral expression vector with central DNA flap TRIPΔU3-EF1α (Sirven et al., 2001, Mol. Ther., 3: 438-448) in order to obtain the plasmids pTRIP-SD/SA-S-CTE and pTRIP-SD/SA-S-WPRE.

SP: signal peptide

TM: transmembrane region

P-CMV: cytomegalovirus immediate/early promoter

P-EF1α: EF1α gene promoter

SD/SA: splice donor and acceptor sites

WPRE: sequences of the "Woodchuck Hepatitis Virus posttranscriptional regulatory element" of the woodchuck hepatitis virus

CTE: sequences of the "constitutive transport element" of the Mason-Pfizer simian retrovirus

LTR: long terminal repeat

ΔU3: LTR deleted for the "promoter/enhancer" sequences

cPPT: "polypurine tract cis-active sequence"

CTS: "central termination sequence"

- Figure 25 shows the Western-blot analysis of the expression of the SARS-CoV S by cell lines transduced with the lentiviral vectors TRIP-SD/SA-S-WPRE and TRIP-SD/SA-S-CTE. Cellular extracts were prepared from established lines FrhK4-S-CTE and FrhK4-S-WPRE after transduction with the lentiviral vectors TRIP-SD/SA-S-CTE and TRIP-SD/SA-S-WPRE respectively. They were separated on an 8% SDS acrylamide gel and analyzed by Western blotting with the aid of an anti-S rabbit polyclonal antibody and an anti-rabbit IgG(H+L)

conjugate coupled to peroxidase. A molecular mass ladder (kDa) is presented in the figure.

T-: control extract of FrhK-4 cells

T+: extract of FrhK-4 cells prepared 24 hours after infection with SARS-CoV at a multiplicity of infection of 3.

- Figure 26 relates to the analysis of the expression of Ssol polypeptide by cell lines transduced with the lentiviral vectors TRIP-SD/SA-Ssol-WPRE and TRIP-SD/SA-Ssol-CTE. The secretion of the Ssol polypeptide was determined in the supernatant of a series of cell clones isolated after transduction of FrhK-4 cells with the lentiviral vectors TRIP-SD/SA-Ssol-WPRE and TRIP-SD/SA-Ssol-CTE. 5 µl of supernatant, diluted 1/2 in loading buffer according to Laemmli, were analyzed by Western blotting, visualized with an anti-FLAG monoclonal antibody (M2, Sigma) and an anti-mouse IgG(H+L) conjugate coupled to peroxidase. T-: supernatant of the parental FRhK-4 line. T+: supernatant of BHK cells infected with a recombinant vaccinia virus expressing the Ssol polypeptide. The solid arrow indicates the Ssol polypeptide, while the empty arrow indicates a cross reaction with a protein of cellular origin.

- Figure 27 shows the results relating to the analysis of the purified Ssol polypeptide

A. 8, 2, 0.5 and 0.125 µg of recombinant Ssol polypeptide purified by anti-FLAG affinity chromatography and gel filtration (G75) were separated on 8% SDS polyacrylamide gel. The Ssol polypeptide and variable quantities of molecular mass markers (MM) were visualized by staining with silver nitrate (Gelcode SilverSNAP stain kit II, Pierce).

B. Standard markers for analysis by SELDI-TOF mass spectrometry

IgG: bovine IgG of MM 147300

ConA: conalbumin of MM 77490

HRP: horseradish peroxidase analyzed as a control and of MM 43240

C. Analysis by mass spectrometry (SELDI-TOF) of the recombinant Ssol polypeptide.

The peaks A and B correspond to the single and double charged Ssol polypeptide.

D. Sequencing of the N-terminal end of the recombinant Ssol polypeptide. 5 Edman degradation cycles in liquid phase were carried out on an ABI494 sequencer (Applied Biosystems).

- Figure 28 illustrates the influence of a splicing signal and of the CTE and WPRE sequences on the efficacy of the gene immunization with the aid of plasmid DNA encoding the SARS-CoV S

A. Groups of 7 BALB/c mice were immunized twice at 4 weeks' interval with the aid of 50 µg of plasmid DNA of pCI, pcDNA-S, pCI-S, pcDNA-N and pCI-HA.

B. Groups of 6 BALB/c mice were immunized twice at 4 weeks' interval with the aid of 2 µg, 10 µg or 50 µg of plasmid DNA of pCI, pCI-S, pCI-S-CTE and pCI-S-WPRE.

The immune sera collected 3 weeks after the second immunization were analyzed by indirect ELISA using a lysate of VeroE6 cells infected with SARS-CoV as antigen. The anti-SARS-CoV antibody titers are calculated as the reciprocal of the dilution producing a specific OD of 0.5 after visualization with an anti-mouse IgG polyclonal antibody coupled to peroxidase (NA931V, Amersham) and TMB (KPL).

- Figure 29 shows the seroneutralization of the infectivity of SARS-CoV with the antibodies induced in

mice after gene immunization with the aid of plasmid DNA encoding SARS-CoV S. Pools of immune sera collected 3 weeks after the second immunization were prepared for each of the groups of experiments described in Figure 28 and evaluated for their capacity to seroneutralize the infectivity of 100 TCID50 of SARS-CoV on FRhK-4 cells. 4 points are produced for each of the 2-fold dilutions tested from 1/20. The seroneutralizing titer is calculated according to the Reed and Munsch method as the reciprocal of the dilution neutralizing the infectivity of 2 wells out of 4.

A. Groups by BALB/c mice immunized twice at 4 weeks' interval with the aid of 50 µg of plasmid DNA of pCI, pcDNA-S, pCI-S, pcDNA-N and pCI-HA. □: preimmune serum. ■: immune serum.

B. Groups of BALB/c mice immunized twice at 4 weeks' interval with the aid of 2 µg, 10 µg or 50 µg of plasmid DNA of pCI, pCI-S, pCI-S-CTE and pCI-S-WPRE.

- Figure 30 illustrates the immunoreactivity of the recombinant Ssol polypeptide toward sera from patients suffering from SARS. The reactivity of sera from patients was analyzed by indirect ELISA test against solid phases prepared with the aid of the purified recombinant Ssol polypeptide. The antibodies from patients reacting with the solid phase at a dilution of 1/400 are visualized with a human anti-IgG(H+L) polyclonal antibody coupled to peroxidase (Amersham NA933V) and TMB plus H202 (KPL). The sera of probable SARS cases are identified by a National Reference Center for Influenza Viruses serial number and by the initials of the patient and the number of days elapsed since the onset of symptoms, where appropriate. The TV sera are control sera from subjects which were collected in France before the SARS epidemic which occurred in 2003.

- Figure 31 shows the induction of antibodies directed against SARS-CoV after immunization with the recombinant Ssol polypeptide. Two groups of 6 mice were immunized at 3 weeks' interval with 10 µg of recombinant Ssol polypeptide (Ssol group) adjuvanted with aluminum hydroxide or, as a control, of adjuvant alone (mock group). Three successive immunizations were performed and the immune sera were collected 3 weeks after each of the three immunizations (IS1, IS2, IS3). The immune sera were analyzed per pool for each of the 2 groups by indirect ELISA using a lysate of VeroE6 cells infected with SARS-CoV as antigen. The anti-SARS-CoV antibody titers are calculated as the reciprocal of the dilution producing a specific OD of 0.5 after visualization with an anti-mouse IgG polyclonal antibody coupled to peroxidase (Amersham) and TMB (KPL).

- Figure 32 presents the nucleotide alignment of the sequences of the synthetic gene 040530 with the sequence of the wild-type gene of the SARS-CoV isolate 031589. I-3059 corresponds to nucleotides 21406-25348 of the SARS-CoV isolate 031589 deposited at the C.N.C.M. under the number I-3059 (SEQ ID NO: 4, plasmid pSARS-S) S-040530 is the sequence of the synthetic gene 040530.

- Figure 33 illustrates the use of a synthetic gene for the expression of the SARS-CoV S. Cellular extracts prepared 48 hours after transfection of VeroE6 cells (A) or 293T cells (B) with the plasmids pCI, pCI-S, pCI-S-CTE, pCI-S-WPRE and pCI-Ssynth were separated on 8% SDS acrylamide gel and analyzed by Western blotting with the aid of an anti-S rabbit polyclonal antibody and an anti-rabbit IgG(H+L) polyclonal antibody coupled to peroxidase (NA934V, Amersham). The Western blot is visualized by luminescence (ECL+, Amersham) and acquisition on a digital imaging device (FluorS,

BioRad). The levels of expression of the S protein were measured by quantifying the 2 predominant bands identified on the image.

- Figure 34 presents a diagram for the construction of recombinant vaccinia viruses VV-TG-S, VV-TG-Ssol, VV-TN-S and VV-TN-Ssol

A. The cDNAs for the S protein and the Ssol polypeptide of SARS-CoV were inserted between the BamH1 and Sma1 sites of the transfer plasmid pTG186 in order to obtain the plasmids pTG-S and pTG-Ssol.

B. The sequences of the synthetic promoter 480 were then substituted for those of the 7.5 promoter by exchange of the Nde1-Pst1 fragments of the plasmids pTG186poly, pTG-S and pTG-Ssol in order to obtain the transfer plasmids pTN480, pTN-S and pTN-Ssol.

C. Sequence of the synthetic promoter 480 as contained between the Nde1 and Pst1 sites of the transfer plasmids of the pTN series. An Asc1 site was inserted in order to facilitate subsequent handling. The restriction sites and the promoter sequence are underlined.

D. The recombinant vaccinia viruses are obtained by double homologous recombination *in vivo* between the TK cassette of the transfer plasmids of the pTG and pTN series and the TK gene of the Copenhagen strain of the vaccinia virus.

SP: signal peptide predicted (aa 1-13) with the software signalP v2.0 (Nielsen et al., 1997, Protein Engineering, 10: 1-6)

TM: transmembrane region predicted (aa 1196-1218) with the software TMHMM v2.0 (Sonnhammer et al., 1998, Proc. of Sixth Int. Conf. on Intelligent Systems for Molecular Biology, pp. 175-182, AAAI Press). It should be noted that the amino acids W1194 and P1195 possibly

form part of the transmembrane region with respective probabilities of 0.13 and 0.42.

TK-L, TK-R: left- and right-hand parts of the vaccinia virus thymidine kinase gene

MCS: multiple cloning site

PE: early promoter

PL: late promoter

PL synth: synthetic late promoter 480

- Figure 35 illustrates the expression of the S protein by recombinant vaccinia viruses, analyzed by Western blotting. Cellular extracts were prepared 18 hours after infection of CV1 cells with the recombinant vaccinia viruses VV-TG, VV-TG-S and VV-TN-S at an M.O.I. of 2 (A). As a control, extracts of VeroE6 cells were prepared 8 hours after infection with SARS-CoV at a multiplicity of infection of 2. Cellular extracts were also prepared 18 hours after infection of CV1 cells with the recombinant vaccinia viruses VV-TG-S, VV-TG-Ssol, VV-TN, VV-TN-S and VV-TN-Ssol (B). They were separated on 8% SDS acrylamide gels and analyzed by Western blotting with the aid of an anti-S rabbit polyclonal antibody and an anti-rabbit IgG(H+L) polyclonal antibody coupled to peroxidase (NA934V, Amersham). "1 µl" and "10 µl" indicates the quantities of cellular extracts deposited on the gel. A molecular mass ladder (kDa) is presented in the figure.

SARS-CoV: extract of VeroE6 cells infected with SARS-CoV

Mock: control extract of noninfected cells

- Figure 36 shows the result of a Western-blot analysis of the secretion of the Ssol polypeptide by the recombinant vaccinia viruses.

A. Supernatants of CV1 cells infected with the recombinant vaccinia virus VV-TN, various clones of the VV-TN-Ssol virus and with the viruses VV-TG-Ssol or

VV-TN-Sflag were harvested 18 hours after infection of CV1 cells at an M.O.I. of 2.

B. Supernatants of 293T, FRhK-4, BHK-21 and CV1 cells infected in duplicate (1.2) with the recombinant vaccinia virus VV-TN-Ssol at an M.O.I. of 2 were harvested 18 hours after infection. The supernatant of CV1 cells infected with the virus VV-TN was also harvested as a control (M).

All the supernatants were separated on 8% SDS acrylamide gel according to Laemmli and analyzed by Western blotting with the aid of an anti-FLAG mouse monoclonal antibody and an anti-mouse IgG(H+L) polyclonal antibody coupled to peroxidase (NA931V, Amersham) (A) or with the aid of an anti-S rabbit polyclonal antibody and an anti-rabbit IgG(H+L) polyclonal antibody coupled to peroxidase (NA934V, Amersham) (B).

A molecular mass ladder (kDa) is presented in the figure.

- Figure 37 shows the analysis of the Ssol polypeptide, purified on SDS polyacrylamide gel

10, 5 and 2 µl of recombinant Ssol polypeptide purified by anti-FLAG affinity chromatography were separated on 4 to 15% gradient SDS polyacrylamide gel. The Ssol polypeptide and variable quantities of molecular mass markers (MM) were visualized by staining with silver nitrate (Gelcode SilverSNAP stain kit II, Pierce).

- Figure 38 illustrates the immunoreactivity of the recombinant Ssol polypeptide produced by the recombinant vaccinia virus VV-TN-Ssol toward sera of patients suffering from SARS. The reactivity of sera from patients was analyzed by indirect ELISA test against solid phases prepared with the aid of the

purified recombinant Ssol polypeptide. The antibodies from patients reacting with the solid phase at a dilution of 1/100 and 1/400 are visualized with a human anti-IgG(H+L) polyclonal antibody coupled to peroxidase (Amersham NA933V) and TMB plus H202 (KPL). The sera of probable SARS cases are identified by a National Reference Center for Influenza Virus serial number and by the initials of the patient and the number of days elapsed since the onset of symptoms, where appropriate. The TV sera are control sera from subjects which were collected in France before the SARS epidemic which occurred in 2003.

- Figure 39 shows the anti-SARS-CoV antibody response in mice after immunization with the recombinant vaccinia viruses. Groups of 7 BALB/c mice were immunized by the i.v. route twice at 4 weeks' interval with 106 pfu of recombinant vaccinia viruses VV-TG, VV-TG-HA, VV-TG-S, VV-TG-Ssol, VV-TN, VV-TN-S, VV-TN-Ssol.

A. Pools of immune sera collected 3 weeks after each of the two immunizations were prepared for each of the groups and were analyzed by indirect ELISA using a lysate of VeroE6 cells infected with SARS-CoV as antigen. The anti-SARS-CoV antibody titers are calculated as the reciprocal of the dilution producing a specific OD of 0.5 after visualization with an anti-mouse IgG polyclonal antibody coupled to peroxidase (NA931V, Amersham) and TMB (KPL).

B. The pools of immune sera were evaluated for their capacity to seroneutralize the infectivity of 100 TCID50 of SARS-CoV on FRhK-4 cells. 4 points are produced for each of the 2-fold dilutions tested from 1/20. The seroneutralizing titer is calculated according to the Reed and Munsch method as the reciprocal of the dilution neutralizing the infectivity of 2 wells out of 4.

- Figure 40 describes the construction of the recombinant viruses MVSchw2-SARS-S and MVSchw2-SARS-Ssol.

A. The measles vector is a complete genome of the Schwarz vaccine strain of the measles virus (MV) into which an additional transcription unit has been introduced (Combredet, 2003, Journal of Virology, 77: 11546-11554). The expression of the additional open reading frames (ORF) is controlled by cis-acting elements necessary for the transcription, for the formation of the cap and for the polyadenylation of the transgene which were copied from the elements present at the N/P junction. 2 different vectors allow the insertion between the P (phosphoprotein) and M (matrix) genes on the one hand and the H (hemagglutinin) and L (polymerase) genes on the other hand.

B. The recombinant genomes MVSchw2-SARS-S and MVSchw2-SARS-Ssol of the measles virus were constructed by inserting the ORFs of the S protein and of the Ssol polypeptide into an additional transcription unit located between the P and M genes of the vector.

The various genes of the measles virus (MV) are indicated: N (nucleoprotein), PVC (V/C phosphoprotein and protein), M (matrix), F (fusion), H (hemagglutinin), L (polymerase). T7 = T7 RNA polymerase promoter, hh = hammerhead ribozyme, T7t = T7 phage RNA polymerase terminator sequence, δ = ribozyme of the hepatitis δ virus, (2), (3) = additional transcription units (ATU).

Size of the MV genome: 15 894 nt.

SP: signal peptide

TM: transmembrane region

FLAG: FLAG tag

- Figure 41 illustrates the expression of the S protein by the recombinant measles viruses, analyzed by Western blotting.

Cytoplasmic extracts were prepared after infection of Vero cells by different passages of the viruses MVSchw2-SARS-S and MVSchw2-SARS-Ssol and the wild-type virus MWSchw as control. Cellular extracts in loading buffer according to Laemmli were also prepared 8 hours after infection of VeroE6 cells with SARS-CoV at a multiplicity of infection of 3. They were separated on 8% SDS acrylamide gel and analyzed by Western blotting with the aid of an anti-S rabbit polyclonal antibody and an anti-rabbit IgG(H+L) polyclonal antibody coupled to peroxidase (NA934V, Amersham).

A molecular mass ladder (kDa) is presented in the figure.

Pn: nth passage of the virus after coculture of 293-3-46 and Vero cells

SARS-CoV: extract of VeroE6 cells infected with SARS-CoV

Mock: control extract of noninfected VeroE6 cells

- Figure 42 shows the expression of the S protein by the recombinant measles viruses, analyzed by immunofluorescence

Vero cells in monolayers on glass slides were infected with the wild-type virus MWSchw (A) or the viruses MVSchw2-SARS-S (B) and MVSchw2-SARS-Ssol (C). When the syncytia have reached 30 to 40% confluence (A., B.) or 90-100% (C), the cells were fixed, permeabilized and labeled with anti-SARS-CoV rabbit polyclonal antibodies and an anti-rabbit IgG(H+L) conjugate coupled to FITC (Jackson).

- Figure 43 illustrates the Western-blot analysis of the immunoreactivity of rabbit sera directed against the peptides E1-12, E53-76 and M2-14. The rabbit 20047 was immunized with the peptide E1-12 coupled to KLH. The rabbits 22234 and 22240 were immunized with the peptide E53-76 coupled to KLH. The rabbits 20013 and 20080 were immunized with the peptide M2-14 coupled to KLH. The immune sera were analyzed by Western blotting with the aid of extracts of cells infected with SARS-CoV (B) or with the aid of extracts of cells infected with a recombinant vaccinia virus expressing the protein E (A) or M (C) of the SARS-CoV 031589 isolate. The immunoblots were visualized with the aid of an anti-rabbit IgG(H+L) conjugate coupled to peroxidase (NA934V, Amersham).

The position of the E and M proteins is indicated by an arrow.

A molecular mass ladder (kDa) is presented in the figure.

It should be understood, however, that these examples are given solely by way of illustration of the subject of the invention, and do not constitute in any manner a limitation thereto.

Example 1: Cloning and sequencing of the genome of the SARS-CoV strain derived from the sample recorded under the number 031589

The RNA of the SARS-CoV strain was extracted from the sample of bronchoalveolar washing recorded under the number 031589, performed on a patient at the Hanoi (Vietnam) French hospital suffering from SARS.

The isolated RNA was used as template to amplify the cDNAs corresponding to the various open reading frames of the genome (ORF1a, ORF1b, ORF-S, ORF-E, ORF-M, ORF-N

(including ORF-13 and ORF-14), ORF3, ORF4, ORF7 to ORF11), and at the noncoding 5′ and 3′ ends. The sequences of the primers and of the probes used for the amplification/detection were defined based on the available SARS-CoV nucleotide sequence.

In the text which follows, the primers and the probes are identified by: the letter S, followed by a letter which indicates the corresponding region of the genome (L for the 5′ end including ORF1a and ORF1b; S, M and N for ORF-S, ORF-M, ORF-N, SE and MN for the corresponding intergene regions), and then optionally by Fn, Rn, with n between 1 and 6 corresponding to the primers used for the nested PCR (F1 + R1 pair for the first amplification, F2 + R2 pair for the second amplication, and the like), and then by /+/ or /-/ corresponding to a sense or antisense primer and finally by the positions of the primers with reference to the Genbank sequence AY27411.3; for the sense and antisense S and N primers and the other sense primers only, when a single position is indicated, it corresponds to that of the 5′ end of a probe or of a primer of about 20 bases; for the antisense primers other than the S and N primers, when a single position is indicated, it corresponds to that of the 3′ end of a probe or of a primer of about 20 bases.

The amplification products thus generated were sequenced with the aid of specific primers in order to determine the complete sequence of the genome of the SARS-CoV strain derived from the sample recorded under the number 031589. These amplification products, with the exception of those corresponding to ORF1a and ORF1b, were then cloned into expression vectors in order to produce the corresponding viral proteins and the antibodies directed against these proteins, in particular by DNA-based immunization.

1. <u>Extraction of the RNAs</u>

The RNAs were extracted with the aid of the *QIamp viral RNA extraction mini* kit (QIAGEN) according to the manufacturer's recommendations. More specifically: 140 µl of the sample and 560 µl of AVL buffer were vigorously mixed for 15 seconds, incubated for 10 minutes at room temperature and then briefly centrifuged at maximum speed. 560 µl of 100% ethanol were added to the supernatant and the mixture thus obtained was very vigorously stirred for 15 sec. 630 µl of the mixture were then deposited on the column.

The column was placed on a 2 ml tube, centrifuged for 1 min at 8000 rpm, and then the remainder of the preceding mixture was deposited on the same column, centrifuged again, for 1 min at 8000 rpm, and the column was transferred over a clean 2 ml tube. Next, 500 µl of AW1 buffer were added to the column, and then the column was centrifuged for 1 min at 8000 rpm and the eluate was discarded. 500 µl of AW2 buffer were added to the column which was then centrifuged for 3 min at 14 000 rpm and transferred onto a 1.5 ml tube. Finally, 60 µl of AVE buffer were added to the column which was incubated for 1 to 2 min at room temperature and then centrifuged for 1 min at 8000 rpm. The eluate corresponding to the purified RNA was recovered and frozen at -20°C.

2. Amplification, sequencing and cloning of the cDNAs

2.1) cDNA encoding the S protein

The RNAs extracted from the sample were subjected to reverse transcription with the aid of random sequence hexameric oligonucleotides (pdN6), so as to produce cDNA fragments.

The sequence encoding the SARS-CoV S glycoprotein was amplified in the form of two overlapping DNA fragments: 5' fragment (SARS-Sa, SEQ ID NO: 5) and 3' fragment

(SARS-Sb, SEQ ID NO: 6), by carrying out two successive amplifications with the aid of nested primers. The amplicons thus obtained were sequenced, cloned into the PCR plasmid vector 2.1-TOPO™ (INVITROGEN), and then the sequence of the cloned cDNAs was determined.

a) Cloning and sequencing of the Sa and Sb fragments

a.1) Synthesis of the cDNA

The reaction mixture containing: RNA (5 µl), H_2O for injection (3.5 µl), 5X reverse transcriptase buffer (4 µl), 5 mM dNTP (2 µl), pdN6 100 µg/ml (4 µl), RNasin 40 IU/µl (0.5 µl) and reverse transcriptase AMV-RT, 10 IU/µl, PROMEGA (1 µl) was incubated in a thermocycler under the following conditions: 45 min at 42°C, 15 min at 55°C, 5 min at 95°C, and then the cDNA obtained was kept at +4°C.

a.2) First PCR amplification

The 5' and 3' ends of the S gene were respectively amplified with the pairs of primers S/F1/+/ 21350-21372 and S/R1/-/ 23518-23498, S/F3/+/ 23258-23277 and S/R3/-/25382-25363. The 50 µl reaction mixture containing: cDNA (2 µl), 50 µM primers (0.5 µl), 10 X buffer (5 µl), 5 mM dNTP (2 µl), Taq Expand High Fidelity, Roche (0.75 µl) and H_2O (39, 75 µl) was amplified in a thermocycler, under the following conditions: an initial step of denaturation at 94°C for 2 min was followed by 40 cycles comprising: a step of denaturation at 94°C for 30 sec, a step of annealing at 55°C for 30 sec and then a step of extension at 72°C for 2 min 30 sec, with 10 sec of additional extension at each cycle, and then a final step of extension at 72°C for 5 min.

a.3) Second PCR amplification

The products of the first PCR amplification (5' and 3' amplicons) were subjected to a second PCR amplification

step (nested PCR) under conditions identical to those of the first amplification, with the pairs of primers S/F2/+/21406-21426 and S/R2/-/23454-23435 and S/F4/+/23322-23341 and S/R4/-/25348-25329, respectively for the 5' amplicon and the 3' amplicon.

a.4) Cloning and sequencing of the Sa and Sb fragments

The Sa (5' end) and Sb (3' end) amplicons thus obtained were purified with the aid of the *QIAquick PCR purification* kit (QIAGEN), following the manufacturer's instructions, and then they were cloned into the vector PCR2.1-TOPO (Invitrogen kit), to give the plasmids called SARS-S1 and SARS-S2.

The DNA of the Sa and Sb clones was isolated and then the corresponding insert was sequenced with the aid of the Big Dye kit, Applied Biosystem® and universal primers M13 forward and M13 reverse, and primers: S/S/+/21867, S/S/+/22353, S/S/+/22811, S/S/+/23754, S/S/+/24207, S/S/+/24699, S/S/+/24348, S/S/-/24209, S/S/-/23630, S/S/-/23038, S/S/-/22454, S/S/-/21815, S/S/-/24784, S/S/+/21556, S/S/+/23130 and S/S/+/24465 following the manufacturer's instructions; the sequences of the Sa and Sb fragments thus obtained correspond to the sequences SEQ ID NO: 5 and SEQ ID NO: 6 in the sequence listing appended as an annex.

The plasmid, called SARS-S1, was deposited under the No. I-3020, on May 12, 2003, at the Collection Nationale de Cultures de Microorganismes, 25 rue du Docteur Roux, 75724 Paris Cedex 15; it contains a 5' fragment of the sequence of the S gene of the SARS-CoV strain derived from the sample recorded under the No. 031589, as defined above, said fragment called Sa corresponding to the nucleotides at positions 21406 to 23454 (SEQ ID NO: 5), with reference to the Genbank sequence AY274119.3 Tor2.

The plasmid, called TOP10F'-SARS-S2, was deposited

under the No. I-3019, on May 12, 2003, at the Collection Nationale de Cultures de Microorganismes, 25 rue du Docteur Roux, 75724 Paris Cedex 15; it contains a 3' fragment of the sequence of the S gene of the SARS-CoV strain derived from the sample recorded under the No. 031589, as defined above, said fragment called Sb corresponding to the nucleotides at positions 23322 to 25348 (SEQ ID NO: 6), with reference to the Genbank sequence accession No. AY274119.3.

b) Cloning and sequencing of the complete cDNA (SARS-S clone of 4 kb)

The complete S cDNA was obtained from the abovementioned clones SARS-S1 and SARS-S2, in the following manner:

1) A PCR amplification reaction was carried out on a SARS-S2 clone in the presence of the abovementioned primer S/R4/-/25348-25329 and of the primer S/S/+/24696-24715: an amplicon of 633 bp was obtained,

2) Another PCR amplification reaction was carried out on another SARS-S2 clone, in the presence of the primers S/F4/+/23322-23341 mentioned above and S/S/-/24803-24784: an amplicon of 1481 bp was obtained.

The amplification reaction was carried out under the conditions as defined above for the amplification of the Sa and Sb fragments, with the exception that 30 amplification cycles comprising a step of denaturation at 94°C for 20 sec and a step of extension at 72°C for 2 min 30 sec were carried out.

3) The 2 amplicons (633 bp and 1481 bp) were purified under the conditions as defined above for the Sa and Sb fragments.

4) Another PCR amplification reaction with the aid of

the abovementioned primers S/F4/+/23322-23341 and S/R4/-/25348-25329 was carried out on the purified amplicons obtained in 3). The amplification reaction was carried out under the conditions as defined above for the amplification of the Sa and Sb fragments, except that 30 amplification cycles were performed.

The 2026 bp amplicon thus obtained was purified, cloned into the vector PCR2.1-TOPO and then sequenced as above, with the aid of the primers as defined above for the Sa and Sb fragments. The clone thus obtained was called clone 3'.

5) The clone SARS-S1 obtained above and the clone 3' were digested with *EcoR I*, the bands of about 2 kb thus obtained were gel purified and then amplified by PCR with the abovementioned primers S/F2/+/21406-21426 and S/R4/-/25348-25329. The amplification reaction was carried out under the conditions as defined above for the amplification of the Sa and Sb fragments, except that 30 amplification cycles were performed. The amplicon of about 4 kb was purified and sequenced. It was then cloned into the vector PCR2.1-TOPO in order to give the plasmid, called SARS-S, and the insert obtained in this plasmid was sequenced as above, with the aid of the primers as defined above for the Sa and Sb fragments. The cDNA sequences of the insert and of the amplicon encoding the S protein correspond respectively to the sequences SEQ ID NO: 4 and SEQ ID NO: 2 in the sequence listing appended as an annex, they encode the S protein (SEQ ID NO: 3).

The sequence of the amplicon corresponding to the cDNA encoding the S protein of the SARS-CoV strain derived from the sample No. 031589 has the following two mutations compared with the corresponding sequences of respectively the Tor2 and Urbani isolates, the positions of the mutations being indicated with reference to the complete sequence of the genome of the

Tor2 isolate (Genbank AY274119.3):
- g/t in position 23220; the alanine codon (gct) in position 577 of the amino acid sequence of the S protein of Tor2 is replaced with a serine codon (tct),
- c/t in position 24872: this mutation does not modify the amino acid sequence of the S protein, and

the plasmid, called SARS-S, was deposited under the No. I-3059, on June 20, 2003, at the Collection Nationale de Cultures de Microorganismes, 25 rue du Docteur Roux, 75724 Paris Cedex 15; it contains the cDNA sequence encoding the S protein of the SARS-CoV strain derived from the sample recorded under the No. 031589, said sequence corresponding to the nucleotides at positions 21406 to 25348 (SEQ ID NO: 4), with reference to the Genbank sequence AY274119.3.

2.2) cDNA encoding the M and E proteins

The RNAs derived from the sample 031589, extracted as above, were subjected to a reverse transcription, combined, during the same step (*Titan One Step RT-PCR®* kit, Roche), with a PCR amplification reaction, with the aid of the pairs of primers:
- S/E/F1/+/26051-26070 and S/E/R1/-/26455-26436 in order to amplify ORF-E, and
- S/M/F1/+/26225-26244 and S/M/R1/-/27148-27129 in order to amplify ORF-M.

A first reaction mixture containing: 8.6 µl of H_2O for injection, 1 µl of dNTP (5 mM), 0.2 µl of each of the primers (50 µM), 1.25 µl of DTT (100 mM) and 0.25 µl of RNAsin (40 IU/µl) was combined with a second reaction mixture containing: 1 µl of RNA, 7 µl of H_2O for injection, 5 µl of 5X RT-PCR buffer and 0.5 µl of enzyme mixture and the combined mixtures were incubated in a thermocycler under the following conditions: 30 min at 42°C, 10 min at 55°C, 2 min at 94°C followed by 40 cycles comprising a step of denaturation at 94°C for 10 sec, a step of annealing at 55°C for 30 sec and

a step of extension at 68°C for 45 sec, with 3 sec increment per cycle and finally a step of terminal extension at 68°C for 7 min.

The amplification products thus obtained (M and E amplicons) were subjected to a second PCR amplification (nested PCR) using the Expand High-Fi® kit, Roche), with the aid of the pairs of primers:
- S/E/F2/+/26082-26101 and S/E/R2/-/26413-26394 for the amplicon E, and
- S/M/F2/+/26330-26350 and S/M/R2/-/27098-27078 for the amplicon M.

The reaction mixture containing: 2 µl of the product of the first PCR, 39.25 µl of H_2O for injection, 5 µl of 10X buffer containing $MgCl_2$, 2 µl of dNTP (5 mM), 0.5 µl of each of the primers (50 µM) and 0.75 µl of enzyme mixture was incubated in a thermocycler under the following conditions: a step of denaturation at 94°C for 2 min was followed by 30 cycles comprising a step of denaturation at 94°C for 15 sec, a step of annealing at 60°C for 30 sec and a step of extension at 72°C for 45 sec, with 3 sec increment per cycle, and finally a step of terminal extension at 72°C for 7 min. The amplification products obtained corresponding to the cDNAs encoding the E and M proteins were sequenced as above, with the aid of the primers: S/E/F2/+/26082 and S/E/R2/-/26394, S/M/F2/+/26330, S/M/R2/-/27078 cited above and the primers S/M/+/26636-26655 and S/M/-/26567-26548. They were then cloned, as above, in order to give the plasmids called SARS-E and SARS-M. The DNA of these clones was then isolated and sequenced with the aid of the universal primers M13 forward and M13 reverse and the primers S/M/+/26636 and S/M/-/26548 mentioned above.

The sequence of the amplicon representing the cDNA encoding the E protein (SEQ ID NO: 13) of the SARS-CoV strain derived from the sample No. 031589 does not

contain differences in relation to the corresponding sequences of the isolates AY274119.3-Tor2 and AY278741-Urbani. The sequence of the E protein of the SARS-CoV 031589 strain corresponds to the sequence SEQ ID NO: 14 in the sequence listing appended as an annex.

The plasmid, called SARS-E, was deposited under the No. I-3046, on May 28, 2003, at the Collection Nationale de Cultures de Microorganismes, 25 rue du Docteur Roux, 75724 Paris Cedex 15; it contains the cDNA sequence encoding the E protein of the SARS-CoV strain derived from the sample recorded under the No. 031589, as defined above, said sequence corresponding to the nucleotides at positions 26082 to 26413 (SEQ ID NO: 15), with reference to the Genbank sequence accession No. AY274119.3.

The sequence of the amplicon representing the cDNA encoding M (SEQ ID NO: 16) from the SARS-CoV strain derived from the sample No. 031589 does not contain differences in relation to the corresponding sequence of the isolate AY274119.3-Tor2. By contrast, at position 26857, the isolate AY278741-Urbani contains a c and the sequence of the SARS-CoV strain derived from the sample recorded under the No. 031589 contains a t. This mutation results in a modification of the amino acid sequence of the corresponding protein: at position 154, a proline (AY278741-Urbani) is changed to serine in the SARS-CoV strain derived from the sample recorded under the No. 031589. The sequence of the M protein of the SARS-CoV strain derived from the sample recorded under the No. 031589 corresponds to the sequence SEQ ID NO: 17 in the sequence listing appended as an annex.

The plasmid, called SARS-M, was deposited under the No. I-3047, on May 28, 2003, at the Collection Nationale de Cultures de Microorganismes, 25 rue du Docteur Roux, 75724 Paris Cedex 15; it contains the cDNA sequence encoding the M protein of the SARS-CoV

strain derived from the sample recorded under the No. 031589, as defined above; said sequence corresponding to the nucleotides at positions 26330 to 27098 (SEQ ID NO: 18), with reference to the Genbank sequence accession No. AY274119.3.

2.3) cDNA corresponding to ORF3, ORF4, ORF7 to ORF11

The same amplification, cloning and sequencing strategy was used to obtain the cDNA fragments corresponding respectively to the following ORFs: ORF3, ORF4, ORF7, ORF8, ORF9, ORF10 and ORF11. The pairs of primers used for the first amplification are:

- ORF3 and ORF4: S/SE/F1/+/25069-25088 and S/SE/R1/-/26300-26281
- ORF7 to ORF11: S/MN/F1/+/26898-26917 and S/MN/R1/-/28287-28266

The pairs of primers used for the second amplification are:

- ORF3 and ORF4: S/SE/F2/+/25110-25129 and S/SE/R2/-/26244-26225
- ORF7 to ORF11: S/MN/F2/+/26977-26996 and S/MN/R2/-/28218-28199

The conditions for the first amplification (RT-PCR) are the following: 45 min at 42°C, 10 min at 55°C, 2 min at 94°C followed by 40 cycles comprising a step of denaturation at 94°C for 15 sec, a step of annealing at 58°C for 30 sec and a step of extension at 68°C for 1 min, with 5 sec increment per cycle and finally a step of terminal extension at 68°C for 7 min.

The conditions for the nested PCR are the following: a step of denaturation at 94°C for 2 min was followed by 40 cycles comprising a step of denaturation at 94°C for 20 sec, a step of annealing at 58°C for 30 sec and a step of extension at 72°C for 50 sec, with 4 sec increment per cycle and finally a step of terminal extension at 72°C for 7 min.

The amplification products obtained corresponding to the cDNAs containing respectively ORF3 and 4 and ORF7 to 11 were sequenced with the aid of the primers: S/SE/+/25363, S/SE/+/25835, S/SE/-/25494, S/SE/-/25875, S/MN/+/27839, S/MN/+/27409, S/MN/-/27836, S/MN/-/27799 and cloned as above for the other ORFs, to give the plasmids called SARS-SE and SARS-MN. The DNA of these clones was isolated and sequenced with the aid of these same primers and of the universal primers M13 sense and M13 antisense.

The sequence of the amplicon representing the cDNA of the region containing OFR3 and ORF4 (SEQ ID NO: 7) of the SARS-CoV strain derived from the sample No. 031589 contains a nucleotide difference in relation to the corresponding sequence of the isolate AY274119-Tor2. This mutation at position 25298 results in a modification of the amino acid sequence of the corresponding protein (ORF3): at position 11, an arginine (AY274119-Tor2) is changed to glycine in the SARS-CoV strain derived from the sample No. 031589. By contrast, no mutation was identified in relation to the corresponding sequence of the isolate AY278741-Urbani. The sequences of ORF3 and 4 of the SARS-CoV strain derived from the sample No. 031589 correspond respectively to the sequences SEQ ID NO: 10 and 12 in the sequence listing appended as an annex.

The plasmid, called SARS-SE, was deposited under the No. I-3126, on November 13, 2003, at the Collection Nationale de Cultures de Microorganismes, 25 rue du Docteur Roux, 75724 Paris Cedex 15; it contains the cDNA corresponding to the region situated between ORF-S and ORF-E and overlapping ORF-E of the SARS-CoV strain derived from the sample recorded under the No. 031589, as defined above, said region corresponding to the nucleotides at positions 25110 to 26244 (SEQ ID NO: 8), with reference to the Genbank sequence accession No. AY274119.3.

The sequence of the amplicon representing the cDNA corresponding to the region containing ORF7 to ORF11 (SEQ ID NO: 19) of the SARS-CoV strain derived from the sample No. 031589 does not contain differences in relation to the corresponding sequences of the isolates AY274119-Tor2 and AY278741-Urbani. The sequences of ORF7 to 11 of the SARS-CoV strain derived from the sample No. 031589 correspond respectively to the sequences SEQ ID NO: 22, 24, 26, 28 and 30 in the sequence listing appended as an annex.

The plasmid, called SARS-MN, was deposited under the No. I-3125, on November 13, 2003, at the Collection Nationale de Cultures de Microorganismes, 25 rue du Docteur Roux, 75724 Paris Cedex 15; it contains the cDNA sequence corresponding to the region situated between ORF-M and ORF-N of the SARS-CoV strain derived from the sample recorded under the No. 031589 and collected in Hanoi, as defined above, said sequence corresponding to the nucleotides at positions 26977 to 28218 (SEQ ID NO: 20), with reference to the Genbank sequence accession No. AY274119.3.

The sequence of the amplicon representing the cDNA corresponding to the region containing ORF7 to ORF11 (SEQ ID NO: 19) of the SARS-CoV strain derived from the sample No. 031589 does not contain differences in relation to the corresponding sequences of the isolates AY274119-Tor2 and AY278741-Urbani. The sequences of ORF7 to 11 of the SARS-CoV strain derived from the sample No. 031589 correspond respectively to the sequences SEQ ID NO: 22, 24, 26, 28 and 30 in the sequence listing appended as an annex.

2.4) cDNA encoding the N protein and including ORF13 and ORF14

The cDNA was synthesized and amplified as described

above for the fragments Sa and Sb. More specifically, the reaction mixture containing: 5 µl of RNA, 5 µl of H_2O for injection, 4 µl of 5X reverse transcriptase buffer, 2 µl of dNTP (5 mM), 2 µl of oligo 20T (5 µM), 0.5 µl of RNasin (40 IU/µl) and 1.5 µl of AMV-RT (10 IU/µl Promega) was incubated in a thermocycler under the following conditions: 45 min at 42°C, 15 min at 55°C, 5 min at 95°C, and it was then kept at +4°C.

A first PCR amplification was performed with the pair of primers S/N/F3/+/28023 and S/N/R3/-/29480.

The reaction mixture as above for the amplification of the S1 and S2 fragments was incubated in a thermocycler, under the following conditions: an initial step of denaturation at 94°C for 2 min was followed by 40 cycles comprising a step of denaturation at 94°C for 20 sec, a step of annealing at 55°C for 30 sec and then a step of extension at 72°C for 1 min 30 sec with 10 sec of additional extension at each cycle, and then a final step of extension at 72°C for 5 min.

The amplicon obtained at the first PCR amplification was subjected to a second PCR amplification step (nested PCR) with the pairs of primer S/N/F4/+/28054 and S/N/R4/-/29430 under conditions identical to those of the first amplification.

The amplification product obtained, corresponding to the cDNA encoding the N protein of the SARS-CoV strain derived from the sample No. 031589, was sequenced with the aid of the primers: S/N/F4/+/28054, S/N/R4/-/29430, S/N/+/28468, S/N/+/28918 and S/N/-/28607 and cloned as above for the other ORFs, to give the plasmid called SARS-N. The DNA of these clones was isolated and sequenced with the aid of the universal primers M13 sense and M13 antisense, and the primers S/N/+/28468, S/N/+/28918 and S/N/-/28607.

The sequence of the amplicon representing the cDNA corresponding to ORF-N and including ORF13 and ORF14 (SEQ ID NO: 36) of the SARS-CoV strain derived from the sample No. 031589 does not contain differences in relation to the corresponding sequences of the isolates AY274119.3-Tor2 and AY278741-Urbani. The sequence of the N protein of the SARS-CoV strain derived from the sample No. 031589 corresponds to the sequence SEQ ID NO: 37 in the sequence listing appended as an annex.

The sequences of ORF13 and 14 of the SARS-CoV strain derived from the sample No. 031589 correspond respectively to the sequences SEQ ID NO: 32 and 34 in the sequence listing appended as an annex.

The plasmid, called SARS-N, was deposited under the No. I-3048, on June 5, 2003, at the Collection Nationale de Cultures de Microorganismes, 25 rue du Docteur Roux, 75724 Paris Cedex 15; it contains the cDNA encoding the N protein of the SARS-CoV strain derived from the sample recorded under the No. 031589, as defined above, said sequence corresponding to the nucleotides at positions 28054 to 29430 (SEQ ID NO: 38), with reference to the Genbank sequence accession No. AY274119.3.

2.5) Noncoding 5' and 3' ends

a) Noncoding 5' end (5'NC)

a_1) Synthesis of the cDNA

The RNAs derived from the sample 031589, extracted as above, were subjected to reverse transcription under the following conditions:

The RNA (15 µl) and the primer S/L/-/443 (3 µl at the concentration of 5 µm) were incubated for 10 min at 75°C.

Next, the 5X reverse transcriptase buffer (6 µl, INVITROGEN), 10 Mm dNTP (1 µl), 0.1 M DTT (3 µl) were added and the mixture was incubated at 50°C for 3 min.

Finally, the reverse transcriptase (3 µl of Superscript®, INVITROGEN) was added to the preceding mixture which was incubated at 50°C for 1 h 30 min and then at 90°C for 2 min.

The cDNA thus obtained was purified with the aid of the *QIAquick PCR purification* kit (QIAGEN), according to the manufacturer's recommendations.

b_1) Terminal transferase reaction (TdT)

The cDNA (10 µl) is incubated for 2 min at 100°C, stored in ice, and the following are then added: H_2O (2.5 µl), 5X TdT buffer (4 µl, AMERSHAM), 5 mM dATP (2 µl) and TdT (1.5 µl, AMERSHAM). The mixture thus obtained is incubated for 45 min at 37°C and then for 2 min at 65°C.

The product obtained is amplified by a first PCR reaction with the aid of the primers: S/L/-225-206 and anchor 14T: 5'-AGATGAATTCGGTACCTTTTTTTTTTTTT-3' (SEQ ID NO: 68). The amplification conditions are the following: an initial step of denaturation at 94°C for 2 min is followed by 10 cycles comprising a step of denaturation at 94°C for 10 sec, a step of annealing at 45°C for 30 sec and then a step of extension at 72°C for 30 sec and then by 30 cycles comprising a step of denaturation at 94°C for 10 sec, a step of annealing at 50°C for 30 sec and then a step of extension at 72°C for 30 sec, and then a final step of extension at 72°C for 5 min.

The product of the first PCR amplification was subjected to a second amplification step with the aid

of the primers: S/L/-/204-185 and anchor 14T mentioned above under conditions identical to those of the first amplification. The amplicon thus obtained was purified, sequenced with the aid of the primer S/L/-/182-163 and it was then cloned as above for the different ORFs, to give the plasmid called SARS-5'NC. The DNA of this clone was isolated and sequenced with the aid of the universal primers M13 sense and M13 antisense and the primer S/L/-/182-163 mentioned above.

The amplicon representing the cDNA corresponding to the 5'NC end of the SARS-CoV strain derived from the sample recorded under the No. 031589 corresponds to the sequence SEQ ID NO: 72 in the sequence listing appended as an annex; this sequence does not contain differences in relation to the corresponding sequences of the isolates AY274119.3-Tor2 and AY278741-Urbani.

The plasmid, called SARS-5'NC, was deposited under the No. I-3124, on November 7, 2003, at the Collection Nationale de Cultures de Microorganismes, 25 rue du Docteur Roux, 75724 Paris Cedex 15; it contains the cDNA corresponding to the noncoding 5' end of the genome of the SARS-CoV strain derived from the sample recorded under the No. 031589, as defined above, said sequence corresponding to the nucleotides at positions 1 to 204 (SEQ ID NO: 39), with reference to the Genbank sequence accession No. AY274119.3.

b) Noncoding 3' end (3'NC)

a_1) Synthesis of the cDNA

The RNAs derived from the sample 031589, extracted as above, were subjected to reverse transcription, according to the following protocol: the reaction mixture containing: RNA (5 µl), H_2O (5 µl), 5X reverse transcriptase buffer (4 µl), 5 mM dNTP (2 µl), 5 µM Oligo 20T (2 µl), 40 U/µl RNasin (0.5 µl) and 10 IU/µl

RT-AMV (1.5 µl, PROMEGA) was incubated in a thermocycler, under the following conditions: 45 min at 42°C, 15 min at 55°C, 5 min at 95°C, and it was then kept at +4°C.

The cDNA obtained was amplified by a first PCR reaction with the aid of the primers S/N/+/28468-28487 and anchor 14T mentioned above. The amplification conditions are the following: an initial step of denaturation at 94°C for 2 min is followed by 10 cycles comprising a step of denaturation at 94°C for 20 sec, a step of annealing at 45°C for 30 sec and then a step of extension at 72°C for 50 sec and then 30 cycles comprising a step of denaturation at 94°C for 20 sec, a step of annealing at 50°C for 30 sec and then a step of extension at 72°C for 50 sec, and then a final step of extension at 72°C for 5 min.

The product of the first PCR amplification was subjected to a second amplification step with the aid of the primers S/N/+/28933-28952 and anchor 14T mentioned above, under conditions identical to those of the first amplification. The amplicon thus obtained was purified, sequenced with the aid of the primer S/N/+/29257-29278 and cloned as above for the different ORFs, to give the plasmid called SARS-3'NC. The DNA of this clone was isolated and sequenced with the aid of the universal primers M13 sense and M13 antisense and the primer S/N/+/29257-29278 mentioned above.

The amplicon representing the cDNA corresponding to the 3'NC end of the SARS-CoV strain derived from the sample recorded under the No. 031589 corresponds to the sequence SEQ ID NO: 73 in the sequence listing appended as an annex; this sequence does not contain differences in relation to the corresponding sequences of the isolates AY274119.3-Tor2 and AY278741-Urbani.

The plasmid called SARS-3'NC was deposited under the

No. I-3123 on November 7, 2003, at the Collection Nationale de Cultures de Microorganismes, 25 rue du Docteur Roux, 75724 Paris Cedex 15; it contains the cDNA sequence corresponding to the noncoding 3' end of the genome of the SARS-CoV strain derived from the sample recorded under the No. 031589, as defined above, said sequence corresponding to that situated between the nucleotide at positions 28933 to 29727 (SEQ ID NO: 40), with reference to the Genbank sequence accession No. AY274119.3, ends with a series of nucleotides a.

2.6) ORF1a and ORF1b

The amplification of the 5' region containing ORF1a and ORF1b of the SARS-CoV genome derived from the sample 031589 was performed by carrying out RT-PCR reactions followed by nested PCRs according to the same principles as those described above for the other ORFs. The amplified fragments overlap over several tenths of bases, thus allowing computer reconstruction of the complete sequence of this part of the genome. On average, the amplified fragments are of two kilobases.

14 overlapping fragments, called L0 to L12, were thus amplified with the aid of the following primers:

Table II: Primers used for the amplification of the 5' region (ORF1a and ORF1b)

REGION AMPLIFIED AND SEQUENCED (does not include the primers)	RT-PCR sense primer	RT-PCR antisense primer	Nested PCR sense primer	Nested PCR antisense primer
L0 50-480	S/L0/F1/+30	S/L0/R1/-481		
L1 231-2240	S/L1/F1/+147	S/L1/R1/-2336	S/L1/F2/+211	S/L1/R2/-2241
L2 2156-4167	S/L2/F1/+2033	S/L2/R1/-4192	S/L2/F2/+2136	S/L2/R2/-4168
L3 3913-5324	S/L3bis/F1/+3850	S/L3bis/R1/-5365	S/L3bis/F2/+3892	S/L3bis/R2/-5325
L4b 4952-6023	S/L4b/F1/+4878	S/L4b/R1/-6061	S/L4b/F2/+4932	S/L4b/R2/-6024
L4 5325-7318	S/L4/F1/+5272	S/L4/R1/-7392	S/L4/F2/+5305	S/L4/R2/-7323
L5 7296-9156	S/L5/F1/+7111	S/L5/R1/-9253	S/L5/F2/+7275	S/L5/R2/-9157
L6 9053-11066	S/L6/F1/+8975	S/L6/R1/-11151	S/L6/F2/+9032	S/L6/R2/-11067
L7 10928-12962	S/L7/F1/+10883	S/L7/R1/-13050	S/L7/F2/+10928	S/L7/R2/-12963
L8 12835-14834	S/L8/F1/+12690	S/L8/R1/-14857	S/L8/F2/+12815	S/L8/R2/-14835
L9 14765-16624	S/L9/F1/+14688	S/L9/R1/-16678	S/L9/F2/+14745	S/L9/R2/-16625
L10 16534-18570	S/L10/F1/+16451	S/L10/R1/-18594	S/L10/F2/+16514	S/L10/R2/-18571
L11 18521-20582	S/L11/F1/+18441	S/L11/R1/-20612	S/L11/F2/+18500	S/L11/R2/-20583
L12 20338-22205.	S/L12/F1/+20279	S/L12/R1/-22229	S/L12/F2/+20319	S/L12/R2/-22206

All the fragments were amplified under the following conditions, except fragment L0 which was amplified as described above for ORF-M:

- RT-PCR: 30 min at 42°C, 15 min at 55°C, 2 min at 94°C, and then the cDNA obtained is amplified under the following conditions: 40 cycles comprising: a step of denaturation at 94°C for 15 sec, a step of annealing at 58°C for 30 sec and then a step of extension at 68°C for 1 min 30 sec, with 5 sec additional extension at each cycle, and then a final step of extension at 68°C for 7 min.

- Nested PCR: An initial step of denaturation at 94°C for 2 min is followed by 35 cycles comprising: a

step of denaturation at 94°C for 15 sec, a step of annealing at 60°C for 30 sec and then a step of extension at 72°C for 1 min 30 sec, with 5 sec of additional extension at each cycle, and then a final step of extension at 72°C for 7 min.

The amplification products were sequenced with the aid of the primers defined in table III below:

Table III: Primers used for the sequencing of the 5' region (ORF1a and ORF1b)

Names	Sequences (SEQ ID NO: 76 to 139)
S/L3/+/4932	5'-CCACACACAGCTTGTGGATA-3'
S/L4/+/6401	5'-CCGAAGTTGTAGGCAATGTC-3'
S/L4/+/6964	5'-TTTGGTGCTCCTTCTTATTG-3'
S/L4/-/6817	5'-CCGGCATCCAAACATAATTT-3'
S/L5/-/7633	5'-TGGTCAGTAGGGTTGATTGG-3'
S/L5/-/8127	5'-CATCCTTTGTGTCAACATCG-3'
S/L5/-/8633	5'-GTCACGAGTGACACCATCCT-3'
S/L5/+/7839	5'-ATGCGACGAGTCTGCTTCTA-3'
S/L5/+/8785	5'-TTCATAGTGCCTGGCTTACC-3'
S/L5/+/8255	5'-ATCTTGGCGCATGTATTGAC-3'
S/L6/-/9422	5'-TGCATTAGCAGCAACAACAT-3'
S/L6/-/9966	5'-TCTGCAGAACAGCAGAAGTG-3'
S/L6/-/10542	5'-CCTGTGCAGTTTGTCTGTCA-3'
S/L6/+/10677	5'-CCTTGTGGCAATGAAGTACA-3'
S/L6/+/10106	5'-ATGTCATTTGCACAGCAGAA-3'
S/L6/+/9571	5'-CTTCAATGGTTTGCCATGTT-3'
S/L7/-/11271	5'-TGCGAGCTGTCATGAGAATA-3'
S/L7/-/11801	5'-AACCGAGAGCAGTACCACAG-3'
S/L7/-/12383	5'-TTTGGCTGCTGTAGTCAATG-3'
S/L7/+/12640	5'-CTACGACAGATGTCCTGTGC-3'
S/L7/+/12088	5'-GAGCAGGCTGTAGCTAATGG-3'
S/L7/+/11551	5'-TTAGGCTATTGTTGCTGCTG-3'
S/L8/-13160	5'-CAGACAACATGAAGCACCAC-3'
S/L8/-/13704	5'-CGCTGACGTGATATATGTGG-3'
S/L8/-14284	5'-TGCACAATGAAGGATACACC-3'
S/L8/+/14453	5'-ACATAGCTCGCGTCTCAGTT-3'
S/L8/+/13968	5'-GGCATTGTAGGCGTACTGAC-3'
S/L8/+/13401	5'-GTTTGCGGTGTAAGTGCAG-3'
S/L9/-15098	5'-TAGTGGCGGCTATTGACTTC-3'
S/L9/-15677	5'-CTAAACCTTGAGCCGCATAG-3'
S/L9/-16247	5'-CATGGTCATAGCAGCACTTG-3'
S/L9/+16323	5'-CCAGGTTGTGATGTCACTGAT-3'
S/L9/+15858	5'-CCTTACCCAGATCCATCAAG-3'
S/L9/+15288	5'-CGCAAACATAACACTTGCTG-3'
S/L10/-16914	5'-AGTGTTGGGTACAAGCCAGT-3'
S/L10/-17466	5'-GTTCCAAGGAACATGTCTGG-3'
S/L10/-18022	5'-AGGTGCCTGTGTAGGATGAA-3'
S/L10/+18245	5'-GGGCTGTCATGCAACTAGAG-3'
S/L10/+17663	5'-TCTTACACGCAATCCTGCTT-3'
S/L10/+17061	5'-TACCCATCTGCTCGCATAGT-3'
S/L11/-/18877	5'-GCAAGCAGAATTAACCCTCA-3'
S/L11/-19396	5'-AGCACCACCTAAATTGCATC-3'
S/L11/-20002	5'-TGGTCCCTTTGAAGGTGTTA-3'
S/L11/+20245	5'-TCGAACACATCGTTTATGGA-3'
S/L11/+/19611	5'-GAAGCACCTGTTTCCATCAT-3'
S/L11/+/19021	5'-ACGATGCTCAGCCATGTAGT-3'
SARS/L1/F3/+800	5'-GAGGTGCAGTCACTCGCTAT-3'
SARS/L1/F4/+1391	5'-CAGAGATTGGACCTGAGCAT-3'

SARS/L1/F5/+1925	5'-CAGCAAACCACTCAATTCCT-3'
SARS/L1/R3/-1674	5'-AAATGATGGCAACCTCTTCA-3'
SARS/L1/R4/-1107	5'-CACGTGGTTGAATGACTTTG-3'
SARS/L1/R5/-520	5'-ATTTCTGCAACCAGCTCAAC-3'
SARS/L2/F3/+2664	5'-CGCATTGTCTCCTGGTTTAC-3'
SARS/L2/F4/+3232	5'-GAGATTGAGCCAGAACCAGA-3'
SARS/L2/F5/+3746	5'-ATGAGCAGGTTGTCATGGAT-3'
SARS/L2/R3/-3579	5'-CTGCCTTAAGAAGCTGGATG-3'
SARS/L2/R4/-2991	5'-TTTCTTCACCAGCATCATCA-3'
SARS/L2/R5/-2529	5'-CACCGTTCTTGAGAACAACC-3'
SARS/L3/F3/+4708	5'-TCTTTGGCTGGCTCTTACAG-3'
SARS/L3/F4/+5305	5'-GCTGGTGATGCTGCTAACTT-3'
SARS/L3/F5/+5822	5'-CCATCAAGCCTGTGTCGTAT-3'
SARS/L3/R3/-5610	5'-CAGGTGGTGCAGACATCATA-3'
SARS/L3/R4/-4988	5'-AACATCAGCACCATCCAAGT-3'
SARS/L3/R5/-4437	5'-ATCGGACACCATAGTCAACG-3'

The sequences of the fragments L0 to L12 of the SARS-CoV strain derived from the sample recorded under the No. 031589 correspond respectively to the sequences SEQ ID NO: 41 to SEQ ID NO: 54 in the sequence listing appended as an annex. Among these sequences, only that corresponding to the fragments L5 contains a nucleotide difference in relation to the corresponding sequence of the isolate AY278741-Urbani. This t/c mutation at position 7919 results in a modification of the amino acid sequence of the corresponding protein, encoded by ORF1a: at position 2552, a valine (gtt codon; AY278741) is changed to alanine (gct codon) in the SARS-CoV strain 031589. By contrast, no mutation was identified in relation to the corresponding sequence of the isolate AY274119.3-Urbani. The other fragments do not exhibit differences in relation to the corresponding sequences of the isolates Tor2 and Urbani.

Example 2: Production and purification of the recombinant N and S proteins of the SARS-CoV strain derived from the sample recorded under the number 031589

The entire N protein and two polypeptide fragments of the S protein of the SARS-CoV strain derived from the sample recorded under the number 031589 were produced in *E. coli*, in the form of fusion proteins comprising

an N- or C-terminal polyhistidine tag. In the two S polypeptides, the N- and C-terminal hydrophobic sequences of the S protein (signal peptide: positions 1 to 13 and transmembrane helix: positions 1196 to 1218) were deleted whereas the β helix (positions 565 to 687) and the two motifs of the coiled-coil type (positions 895 to 980 and 1155 to 1186) of the S protein were preserved. These two polypeptides consist of: a long fragment (S_L) corresponding to positions 14 to 1193 of the amino acid sequence of the S protein and a short fragment (S_C) corresponding to positions 475 to 1193 of the amino acid sequence of the S protein.

1) Cloning of the cDNAs N, S_L and S_C into the expression vectors pIVEX2.3 and pIVEX2.4

The cDNAs corresponding to the N protein and to the S_L and S_C fragments were amplified by PCR under standard conditions, with the aid of the DNA polymerase Platinium Pfx® (INVITROGEN). The plasmids SRAS-N and SRAS-S were used as template and the following oligonucleotides as primers:

5'-CCCATATGTCTGATAATGGACCCCAATCAAAC-3' (N sense, SEQ ID NO: 55)

5'-CCCCCGGGTGCCTGAGTTGAATCAGCAGAAGC-3' (N antisense, SEQ ID NO: 56)

5'-CCCATATGAGTGACCTTGACCGGTGCACCAC-3' (S_C sense, SEQ ID NO: 57)

5'-CCCATATGAAACCTTGCACCCCACCTGCTC-3' (S_L sense, SEQ ID NO: 58)

5'-CCCCCGGGTTTAATATATTGCTCATATTTTCCC-3' (S_C and S_L antisense, SEQ ID NO: 29).

The sense primers introduce an *NdeI* site (underlined) while the antisense primers introduce an *XmaI* or *SmaI* site (underlined). The 3 amplification products were column purified (*QIAquick PCR Purification* kit, QIAGEN) and cloned into an appropriate vector. The plasmid DNA purified from the 3 constructs (*QIAFilter Midi Plasmid*

kit, QIAGEN) was verified by sequencing and digested with the enzymes *NdeI* and *XmaI*. The 3 fragments corresponding to the cDNAs N, S_L and S_C were purified on agarose gel and then inserted into the plasmids pIVEX2.3MCS (C-terminal polyhistidine tag) and pIVEX2.4d (N-terminal polyhistidine tag) digested beforehand with the same enzymes. After verification of the constructs, the 6 expression vectors thus obtained (pIV2.3N, pIV2.3S_C, pIV2.3S_L, pIV2.4N, pIV2.4S_C also called pIV2.4S_1, pIV2.4S_L) were then used, on the one hand to test the expression of the proteins *in vitro*, and on the other hand to transform the bacterial strain BL21(DE3)pDIA17 (NOVAGEN). These constructs encode proteins whose expected molecular mass is the following: pIV2.3N (47174 Da), pIV2.3S_C (82897 Da), pIV2.3S_L (132056 Da), pIV2.4N (48996 Da), pIV2.4S_1 (81076 Da) and pIV2.4S_L (133877 Da). Bacteria transformed with pIV2.3N were deposited at the CNCM on October 23, 2003, under the number I-3117, and bacteria transformed with pIV2.4S_1 were deposited at the CNCM on October 23, 2003, under the number I-3118.

2) Analysis of the expression of the recombinant proteins *in vitro* and *in vivo*

The expression of recombinant proteins from the 6 recombinant vectors was tested, in a first instance, in a system *in vitro* (RTS100, Roche). The proteins produced *in vitro*, after incubation of the recombinant vectors pIVEX for 4 h at 30°C, in the RTS100 system, were analyzed by Western blotting with the aid of an anti-$(his)_6$ antibody coupled to peroxidase. The result of expression *in vitro* (figure 1) shows that only the N protein is expressed in large quantities, regardless of the position, N- or C-terminal, of the polyhistidine tag. In a second step, the expression of the N and S proteins was tested *in vivo* at 30°C in LB medium in the presence or in the absence of inducer (1 mM IPTG). The N protein is very well produced in this bacterial

The sequences of the fragments L0 to L12 of the SARS-CoV strain derived from the sample recorded under the No. 031589 correspond respectively to the sequences SEQ ID NO: 41 to SEQ ID NO: 54 in the sequence listing appended as an annex. Among these sequences, only that corresponding to the fragments L5 contains a nucleotide difference in relation to the corresponding sequence of the isolate AY278741-Urbani. This t/c mutation at position 7919 results in a modification of the amino acid sequence of the corresponding protein, encoded by ORF1a: at position 2552, a valine (gtt codon; AY278741) is changed to alanine (gct codon) in the SARS-CoV strain 031589. By contrast, no mutation was identified in relation to the corresponding sequence of the isolate AY274119.3-Urbani. The other fragments do not exhibit differences in relation to the corresponding sequences of the isolates Tor2 and Urbani.

Example 2: Production and purification of the recombinant N and S proteins of the SARS-CoV strain derived from the sample recorded under the number 031589

The entire N protein and two polypeptide fragments of the S protein of the SARS-CoV strain derived from the sample recorded under the number 031589 were produced in *E. coli*, in the form of fusion proteins comprising an N- or C-terminal polyhistidine tag. In the two S polypeptides, the N- and C-terminal hydrophobic sequences of the S protein (signal peptide: positions 1 to 13 and transmembrane helix: positions 1196 to 1218) were deleted whereas the β helix (positions 565 to 687) and the two motifs of the coiled-coil type (positions 895 to 980 and 1155 to 1186) of the S protein were preserved. These two polypeptides consist of: a long fragment (S_L) corresponding to positions 14 to 1193 of the amino acid sequence of the S protein and a short

fragment (S_C) corresponding to positions 475 to 1193 of the amino acid sequence of the S protein.

1) Cloning of the cDNAs N, S_L and S_C into the expression vectors pIVEX2.3 and pIVEX2.4

The cDNAs corresponding to the N protein and to the S_L and S_C fragments were amplified by PCR under standard conditions, with the aid of the DNA polymerase Platinium Pfx® (INVITROGEN). The plasmids SRAS-N and SRAS-S were used as template and the following oligonucleotides as primers:

5'-CCCATATGTCTGATAATGGACCCCAATCAAAC-3' (N sense, SEQ ID NO: 55)

5'-CCCCCGGGTGCCTGAGTTGAATCAGCAGAAGC-3' (N antisense, SEQ ID NO: 56)

5'-CCCATATGAGTGACCTTGACCGGTGCACCAC-3' (S_C sense, SEQ ID NO: 57)

5'-CCCATATGAAACCTTGCACCCCACCTGCTC-3' (S_L sense, SEQ ID NO: 58)

5'-CCCCCGGGTTTAATATATTGCTCATATTTTCCC-3' (S_C and S_L antisense, SEQ ID NO: 29).

The sense primers introduce an *NdeI* site (underlined) while the antisense primers introduce an *XmaI* or *SmaI* site (underlined). The 3 amplification products were column purified (*QIAquick PCR Purification* kit, QIAGEN) and cloned into an appropriate vector. The plasmid DNA purified from the 3 constructs (*QIAFilter Midi Plasmid* kit, QIAGEN) was verified by sequencing and digested with the enzymes *NdeI* and *XmaI*. The 3 fragments corresponding to the cDNAs N, S_L and S_C were purified on agarose gel and then inserted into the plasmids pIVEX2.3MCS (C-terminal polyhistidine tag) and pIVEX2.4d (N-terminal polyhistidine tag) digested beforehand with the same enzymes. After verification of the constructs, the 6 expression vectors thus obtained (pIV2.3N, pIV2.3S_C, pIV2.3S_L, pIV2.4N, pIV2.4S_C also

called pIV2.4S_1, pIV2.4S_L) were then used, on the one hand to test the expression of the proteins *in vitro*, and on the other hand to transform the bacterial strain BL21(DE3)pDIA17 (NOVAGEN). These constructs encode proteins whose expected molecular mass is the following: pIV2.3N (47174 Da), pIV2.3S_C (82897 Da), pIV2.3S_L (132056 Da), pIV2.4N (48996 Da), pIV2.4S_1 (81076 Da) and pIV2.4S_L (133877 Da). Bacteria transformed with pIV2.3N were deposited at the CNCM on October 23, 2003, under the number I-3117, and bacteria transformed with pIV2.4S_1 were deposited at the CNCM on October 23, 2003, under the number I-3118.

2) Analysis of the expression of the recombinant proteins *in vitro* and *in vivo*

The expression of recombinant proteins from the 6 recombinant vectors was tested, in a first instance, in a system *in vitro* (RTS100, Roche). The proteins produced *in vitro*, after incubation of the recombinant vectors pIVEX for 4 h at 30°C, in the RTS100 system, were analyzed by Western blotting with the aid of an anti-$(his)_6$ antibody coupled to peroxidase. The result of expression *in vitro* (figure 1) shows that only the N protein is expressed in large quantities, regardless of the position, N- or C-terminal, of the polyhistidine tag. In a second step, the expression of the N and S proteins was tested *in vivo* at 30°C in LB medium in the presence or in the absence of inducer (1 mM IPTG). The N protein is very well produced in this bacterial system (figure 2) and is found mainly in a soluble fraction after lysis of the bacteria. By contrast, the long version of S (S_L) is very weakly produced and is completely insoluble (figure 3). The short version (S_C) also exhibits a very weak solubility, but an expression level that is much higher than that of the long version. Moreover, the construct S_C fused with a polyhistidine tag at the C-terminal position has a

smaller size than that expected. An immunodetection experiment with an anti-polyhistidine antibody has shown that this construct was incomplete. In conclusion, the two constructs, pIV2.3N and pIV2.4S_1, which express respectively the entire N protein fused with the C-terminal polyhistidine tag and the short S protein fused with the N-terminal polyhistidine tag, were selected in order to produce the two proteins in a large quantity so as to purify them. The plasmids pIV2.3N and pIV2.4S_1 were deposited respectively under the No. I-3117 and I-3118 at the CNCM, 25 rue du Docteur Roux, 75724 PARIS 15, on October 23, 2003.

3) Analysis of the antigenic activity of the recombinant proteins

The antigenic activity of the N, S_L and S_C proteins was tested by Western blotting with the aid of two serum samples, obtained from the same patient infected with SARS-CoV, collected 8 days (M12) and 29 days (M13) after the onset of the SARS symptoms. The experimental protocol is as described in example 3. The results illustrated by figure 4 show (i) the seroconversion of the patient, and (ii) that the N protein possesses a higher antigenic reactivity than the short S protein.

4) Purification of the N protein from pIV2.3N

Several experiments for purifying the N protein, produced from the vector pIV2.3N, were carried out according to the following protocol. The bacteria BL21(DE3)pDIA17, transformed with the expression vector pIV2.3N, were cultured at 30°C in 1 liter of culture medium containing 0.1 mg/ml of ampicillin, and induced with 1 mM IPTG when the cell density equivalent to A_{600} = 0.8 is reached (about 3 hours). After 2 hours of culture in the presence of inducer, the cells were recovered by centrifugation (10 min at 5000 rpm),

resuspended in the lysis buffer (50 mM NaH_2PO_4, 0.3 M NaCl, 20 mM imidazole, pH 8, containing the mixture of protease inhibitors *Complete®*, Roche), and lysed with the French press (12 000 psi). After centrifugation of the bacterial lysate (15 min at 12 000 rpm), the supernatant (50 ml) was deposited at a flow rate of 1 ml/min on a metal chelation column (15 ml) (Ni-NTA superflow, Qiagen), equilibrated with the lysis buffer. After washing the column with 200 ml of lysis buffer, the N protein was eluted with an imidazole gradient (20 → 250 mM) in 10 column volumes. The fractions containing the N protein were assembled and analyzed by polyacrylamide gel electrophoresis under denaturing conditions followed by staining with Coomassie blue. The results illustrated by figure 5 show that the protocol used makes it possible to purify the N protein with a very satisfactory homogeneity (95%) and a mean yield of 15 mg of protein per liter of culture.

5) Purification of the S_C protein from pIV2.4S_C (pIV2.4S_1)

The protocol followed for purifying the short S protein is very different from that described above because the protein is highly aggregated in the bacterial system (inclusion bodies). The bacteria BL21(DE3)pDIA17, transformed with the expression vector pIV2.4S_1, were cultured at 30°C in 1 liter of culture medium containing 0.1 mg/ml of ampicillin, and induced with 1 mM IPTG when the cell density equivalent to A_{600} = 0.8 is reached (about 3 hours). After 2 hours of culture in the presence of inducer, the cells were recovered by centrifugation (10 min at 5000 rpm), resuspended in the lysis buffer (0.1 M Tris-HCl, 1 mM EDTA, pH 7.5), and lysed with the French press (1200 psi). After centrifugation of the bacterial lysate (15 min at 12 000 rpm), the pellet was resuspended in 25 ml of lysis buffer containing 2% Triton X100 and 10 mM

β-mercaptoethanol, and then centrifuged for 20 min at 12 000 rpm. The pellet was resuspended in 10 mM Tris-HCl buffer containing 7 M urea, and gently stirred for 30 min at room temperature. This final washing of the inclusion bodies with 7 M urea is necessary in order to remove most of the *E. coli* membrane proteins which co-sediment with the aggregated S_C protein. After a final centrifugation for 20 min at 12 000 rpm, the final pellet is resuspended in the 10 mM Tris-HCl buffer. The electrophoretic analysis of this preparation (figure 6) shows that the short S protein may be purified with a satisfactory homogeneity (about 90%) from the inclusion bodies (insoluble extract).

Example 3: Immunodominance of the N protein

The reactivity of the antibodies present in the serum of patients suffering from atypical pneumopathy caused by the SARS-associated coronavirus (SARS-CoV), toward the various proteins of this virus, was analyzed by Western blotting under the conditions described below.

1) Materials

a) Lysate of cells infected with SARS-CoV

Vero E6 cells (2×10^6) were infected with SARS-CoV (isolate recorded under the number FFM/MA104) at a multiplicity of infection (M.O.I.) of 10^{-1} or 10^{-2} and then incubated in DMEM medium containing 2% FCS, at 35°C in an atmosphere containing 5% CO_2. 48 hours later, the cellular lawn was washed with PBS and then lysed with 500 µl of loading buffer prepared according to Laemmli and containing β-mercaptoethanol. The samples were then boiled for 10 minutes and then sonicated for 3 times 20 seconds.

b) Antibodies

b_1) Serum from a patient suffering from atypical pneumopathy

The serum designated by a reference at the National Reference Center for Influenza Viruses (Northern region) under the No. 20033168 is that from a French patient suffering from atypical pneumopathy caused by SARS-CoV collected on day 38 after the onset of the symptoms; the diagnosis of SARS-CoV infection was performed by nested RT-PCR and quantitative PCR.

b_2) Monospecific rabbit polyclonal sera directed against the N protein or the S protein

The sera are those produced from the recombinant N and S_C proteins (example 2), according to the immunization protocol described in example 4; they are the rabbit P13097 serum (anti-N serum) and the rabbit P11135 serum (anti-S serum).

2) Method

20 µl of lysate of cells infected with SARS-CoV at M.O.I. values of 10^{-1} and 10^{-2} and, as a control, 20 µl of a lysate of noninfected cells (mock) were separated on 10% SDS polyacrylamide gel and then transferred onto a nitrocellulose membrane. After blocking in a solution of PBS/5% milk/0.1% Tween and washing in PBS/0.1% Tween, this membrane was hybridized overnight at 4°C with: (i) the immune serum No. 20033168 diluted 1/300, 1/1000 and 1/3000 in the buffer PBS/1% BSA/0.1% Tween, (ii) the rabbit P13097 serum (anti-N serum) diluted 1/50 000 in the same buffer and (iii) the rabbit P11135 serum (anti-S serum) diluted 1/10 000 in the same buffer. After washing in PBS/Tween, a secondary hybridization was performed with the aid of either sheep polyclonal antibodies directed against the heavy

and light chains of human G immunoglobulins and coupled with peroxidase (NA933V, Amersham), or of donkey polyclonal antibodies directed against the heavy and light chains of the rabbit G immunoglobulins and coupled with peroxidase (NA934V, Amersham). The bound antibodies were visualized with the aid of the ECL+ kit (Amersham) and of Hyperfilm MP autoradiography films (Amersham). A molecular mass ladder (kDa) is presented in the figure.

3) Results

Figure 7 shows that three polypeptides of apparent molecular mass 35, 55 and 200 kDa are specifically detected in the extracts of cells infected with SARS-CoV.

In order to identify these polypeptides, two other immunoblots (figure 8) were prepared on the same samples and under the same conditions with rabbit polyclonal antibodies specific for the nucleoprotein N (rabbit P13097, figure 8A) and for the spicule protein S (rabbit P11135, figure 8B). This experiment shows that the 200 kDa polypeptide corresponds to the SARS-CoV spicule glycoprotein S, that the 55 kDa polypeptide corresponds to the nucleoprotein N while the 35 kDa polypeptide probably represents a truncated or degraded form of N.

The data presented in figure 7 therefore show that the serum 20033168 strongly reacts with N and a lot more weakly with the SARS-CoV S since the 35 and 55 kDa polypeptides are visualized in the form of intense bands for 1/300, 1/1000 and 1/3000 dilutions of the immunoserum whereas the 200 kDa polypeptide is only weakly visualized for a dilution of 1/300. It is also possible to note that no other SARS-CoV polypeptide is detected for dilutions greater than 1/300 of the serum

20033168.

This experiment indicates that the antibody response specific for the SARS-CoV N dominates the antibody responses specific for the other SARS-CoV polypeptides and in particular the antibody response directed against the S glycoprotein. It indicates an immunodominance of the nucleoprotein N during human infections with SARS-CoV.

Example 4: Preparation of monospecific polyclonal antibodies directed against the SARS-associated coronavirus (SARS-CoV) N and S proteins

1) Materials and method

Three rabbits (P13097, P13081, P13031) were immunized with the purified recombinant polypeptide corresponding to the entire nucleoprotein (N), prepared according to the protocol described in example 2. After a first injection of 0.35 mg per rabbit of protein emulsified in complete Freund's adjuvant (intradermal route), the animals received 3 booster injections at 3 and then 4 weeks' interval, of 0.35 mg of recombinant protein emulsified in incomplete Freund's adjuvant.

Three rabbits (P11135, P13042, P14001) were immunized with the recombinant polypeptide corresponding to the short fragment of the S protein (S_C) produced as described in example 2. As this polypeptide is found mainly in the form of inclusion bodies in the bacterial cytoplasm, the animals received 4 intradermal injections at 3-4 weeks' interval of a preparation of inclusion bodies corresponding to 0.5 mg of recombinant protein emulsified in incomplete Freund's adjuvant. The first 3 injections were made with a preparation of inclusion bodies prepared according to the protocol described in example 2, while the fourth injection was

made with a preparation of inclusion bodies which were prepared according to the protocol described in example 2 and then purified on sucrose gradient and washed in 2% Triton X100.

For each rabbit, a preimmune (p.i.) serum was prepared before the first immunization and an immune serum (I.S.) 5 weeks after the fourth immunization.

In a first instance, the reactivity of the sera was analyzed by ELISA test on preparations of recombinant proteins similar to those used for the immunizations; the ELISA tests were carried out according to the protocol and with the reagents as described in example 6.

In a second instance, the reactivity of the sera was analyzed by preparing an immunoblot (Western blot) of a lysate of cells infected with SARS-CoV, according to the protocol as described in example 3.

2) Results

The ELISA tests (figure 9) demonstrate that the preparations of recombinant N protein and of inclusion bodies of the short fragment of the S protein (S_C) are immunogenic in animals and that the titer of the immune sera is high (more than 1/25 000).

The immunoblot (figure 8) shows that the rabbit P13097 immune serum recognizes two polypeptides present in the lysates of cells infected with SARS-CoV: a polypeptide whose apparent molecular mass (50-55 kDa based on experiments) is compatible with that of the nucleo-protein N (422 residues, predicted molecular mass of 46 kDa) and a polypeptide of 35 kDa, which probably represents a truncated or degraded form of N.

This experiment also shows that the rabbit P11135 serum mainly recognizes a polypeptide whose apparent molecular mass (180-220 kDa based on experiments) is compatible with a glycosylated form of S (1255 residues, nonglycosylated polypeptide chain of 139 kDa), as well as lighter polypeptides, which probably represent truncated and/or nonglycosylated forms of S.

In conclusion, all these experiments demonstrate that the recombinant polypeptides expressed in *E. coli* and corresponding to the SARS-CoV N and S proteins make it possible to induce, in animals, polyclonal antibodies capable of recognizing the native forms of these proteins.

Example 5: Preparation of monospecific polyclonal antibodies directed against the SARS-associated coronavirus (SARS-CoV) M and E proteins

1) Analysis of the structure of the M and E proteins

a) E protein

The structure of the SARS-CoV E protein (76 amino acids) was analyzed *in silico*, with the aid of various software packages such as signalP v1.1, NetNGlyc 1.0, THMM 1.0 and 2.0 (Krogh et al., 2001, J. Mol. Biol., 305(3):567-580) or alternatively TOPPRED (von Heijne, 1992, J. Mol. Biol. 225, 487-494). The analysis shows that this nonglycosylated polypeptide is a type 1 membrane protein, containing a single transmembrane helix (aa 12-34 according to THMM), and in which the majority of the hydrophilic domain (42 residues) is located at the C-terminal end and probably inside the viral particle (endodomain). It is possible to note an inversion in the topology predicted by versions 1.0 (N-ter is external) and 2.0 (N-ter is internal) of the

THMM software, but that other algorithms, in particular TOPPRED and THUMBUP (Zhou et Zhou, 2003, Protein Science 12:1547-1555) confirm an external location of the N-terminal end of E.

b) M protein

A similar analysis carried out on the SARS-CoV M protein (221 amino acids) shows that this polypeptide does not possess a signal peptide (according to the software signalP v1.1) but three transmembrane domains (residues 15-37, 50-72, 77-99 according to THMM2.0) and a large hydrophilic domain (aa 100-221) located inside the viral particle (endodomain). It is probably glycosylated on the asparagine at position 4 (according to NetNGlyc 1.0).

Thus, in agreement with the experimental data known for the other coronaviruses, it is remarkable that the two M and E proteins exhibit endodomains corresponding to the majority of the polypeptides and of the ectodomains that are very small in size.

- The ectodomain of E probably corresponds to residues 1 to 11 or 1 to 12 of the protein: MYSFVSEETGT(L), SEQ ID NO: 70. Indeed, the probability associated with the transmembrane location of residue 12 is intermediate (0.56 according to THMM 2.0).

- The ectodomain of M probably corresponds to residues 2 to 14 of the protein: ADNGTITVEELKQ, SEQ ID NO: 69. Indeed, the N-terminal methionine of M is very probably cleaved from the mature polypeptide because the residue at position 2 is an alanine (Varshavsky, 1996, 93:12142-12149).

Moreover, the analysis of the hydrophobicity (Kyte & Doolittle, Hopp & Woods) of the E protein demonstrates

that the C-terminal end of the endodomain of E is hydrophilic and therefore probably exposed at the surface of this domain. Thus, a synthetic peptide corresponding to this end is a good immunogenic candidate for inducing, in animals, antibodies directed against the endodomain of E. Consequently, a peptide corresponding to 24 C-terminal residues of E was synthesized.

2) Preparation of antibodies directed against the ectodomain of the M and E proteins and the endodomain of the E protein

The peptides M2-14 (ADNGTITVEELKQ, SEQ ID NO: 69), E1-12 (MYSFVSEETGTL, SEQ ID NO: 70) and E53-76 (KPTVYVYSRV KNLNSSEGVP DLLV, SEQ ID NO: 71) were synthesized by Neosystem. They were coupled with KLH (*Keyhole Limpet Hemocyanin*) with the aid of MBS (m-maleimido-benzoyl-N-hydroxysuccinimide ester) via a cysteine added during the synthesis either at the N-terminus of the peptide (case for E53-76) or at the C-terminus (case of M2-14 and E1-12).

Two rabbits were immunized with each of the conjugates, according to the following immunization protocol: after a first injection of 0.5 mg of peptide coupled with KLH and emulsified in complete Freund's adjuvant (intradermal route), the animals receive 2 to 4 booster injections at 3 or 4 weeks' interval of 0.25 mg of peptide coupled to KLH and emulsified in incomplete Freund's adjuvant.

For each rabbit, a preimmune (p.i.) serum was prepared before the first immunization and an immune serum (I.S.) is prepared 3 to 5 weeks after the booster injections.

The reactivity of the sera was analyzed by Western

blotting with the aid of extracts of cells infected with SARS-CoV (figure 43B) or with the aid of extracts of cells infected with a recombinant vaccinia virus expressing the protein E (VV-TG-E, figure 43A) or M (VV-TN-M, figure 43C) of the SARS-CoV 031589 isolate.

The immune sera of the rabbits 22234 and 22240, immunized with the conjugate KLH-E53-76, recognize a polypeptide of about 9 to 10 kD, which is present in the extracts of cells infected with SARS-CoV but absent from the extracts of noninfected cells (figure 43B). The apparent mass of this polypeptide is compatible with the predicted mass of the E protein, which is 8.4 kD. Similarly, the immune serum of the rabbit 20047, immunized with the conjugate KLH-E1-12, recognizes a polypeptide present in the extracts of cells infected with the VV-TG-E virus, whose apparent molar mass is compatible with that of the E protein (figure 43A).

The immune serum of the rabbits 20013 and 20080, immunized with the conjugate KLH-M2-14, recognizes a polypeptide present in the extracts of cells infected with the VV-TN-M virus (figure 43C), whose apparent molar mass (about 18 kD) is compatible with that of the glycoprotein M, which is 25.1 kD and has a high isoelectric point (9.1 for the naked polypeptide).

These results demonstrate that the peptides E1-12 and E53-76, on the one hand, and the peptide M2-14, on the other hand, make it possible to induce, in animals, polyclonal antibodies capable of recognizing the native forms of the SARS-CoV E and M proteins, respectively.

Example 6: Analysis of the ELISA reactivity of the recombinant N protein toward sera from patients suffering from SARS

1) Materials

The antigen used to prepare the solid phases is the purified recombinant nucleoprotein N prepared according to the protocol described in example 2.

The sera to be tested (table IV) were chosen on the basis of the results of analysis of their reactivity by immunofluorescence (IF-SARS titer), toward cells infected with SARS-CoV.

Table IV: Sera tested by ELISA

Reference	Serum No.	Type of serum	Date of the serum***	IF-SARS titer
3050	A	Control	na*	nt**
3048	B	Control	na	nt
033168	D	Patient 1-SARS	04/27/03 (D38)	320
033397	E	Patient-1 SARS	05/11/05 (D52)	320
032632	F	Patient-2 SARS	03/21/03 (D17)	2500
032791	G	Patient-3 SARS	04/04/03 (D3)	<40
033258	H	Patient-3 SARS	04/28/03 (D27)	160

*na: not applicable. **nt: not tested. *** the dates indicated correspond to the number of days after the onset of the SARS symptoms.

2) Method

The N protein (100 µl) diluted at various concentrations in 0.1 M carbonate buffer, pH 9.6 (1, 2 or 4 µg/ml) is distributed into the wells of ELISA plates, and then the plates are incubated overnight at laboratory temperature. The plates are washed with PBS-Tween buffer saturated with PBS-skimmed milk-sucrose (5%) buffer. The test sera (100 µl), diluted beforehand (1/50, 1/100, 1/200, 1/400, 1/800, 1/1600 and 1/3200) are added and then the plates are incubated for 1 h at 37°C. After 3 washings, the peroxidase-

labeled anti-human IgG conjugate (reference 209-035-098, JACKSON) diluted 1/18 000 is added and then the plates are incubated for 1 h at 37°C. After 4 washings, the chromogen (TMB) and the substrate (H_2O_2) are added and the plates are incubated for 30 min at room temperature, protected from light. The reaction is then stopped and then the absorbance at 450 nm is measured with the aid of an automated reader.

3) Results

The ELISA tests (figure 10) demonstrate that the recombinant N protein preparation is specifically recognized by the antibodies of sera from patients suffering from SARS collected in the late phase of the infection (≥ 17 days after the onset of the symptoms) whereas it is not significantly recognized by the antibodies of a patient's serum collected in the early phase of the infection (3 days after the onset of the symptoms) or by control sera from subjects not suffering from SARS.

Example 7: ELISA tests prepared for a very specific and sensitive detection of a SARS-associated coronavirus infection, from sera of patients

1) Indirect ELISA IgG test

a) Reagents

Preparation of the plates

The plates are sensitized with a solution of N protein at 2 µg/ml in a 10 mM PBS buffer, pH 7.2, phenol red at 0.25 ml/l. 100 µl of solution are deposited in the wells and left to incubate at room temperature overnight. Saturation is obtained by prewashing in 10 mM PBS/0.1% Tween buffer, followed by washing with a

saturation solution PBS, 25% milk/sucrose.

Diluent sera
Buffer 0.48 g/l TRIS, 10 mM PBS, 3.7 g/l EDTA, 15% v/v milk, pH 6.7

Diluent conjugate
Citrate buffer (15 g/l), 0.5% Tween, 25% bovine serum, 12% NaCl, 6% v/v skimmed milk pH 6.5

Conjugate
50X anti-human IgG conjugate, marketed by Bio-Rad: Platelia H. pylori kit ref 72778

Other solutions:
Washing solution R2, solutions for visualizing with TMB R8 diluent, R9 chromogen, R10 stopping solution: reagents marketed by Bio-Rad (e.g.: *Platelia pylori* kit, ref 72778)

b) Procedure

Dilute the sera **1/200** in the sample diluent

Distribute 100 µl/well

Incubation 1 h at 37°C

3 washings in 10× WASHING solution R2 diluted beforehand 10-fold in demineralized water (i.e., 1X washing solution)

Distribute 100 µl of conjugate (50× conjugate to be diluted immediately before use in the diluent conjugate provided)

Incubation 1 h at 37°C

4 washings in 1X washing solution

Distribute 200 µl/well of visualization solution (to be diluted immediately before use e.g.: 1 ml of R9 in 10 ml of R8)

Incubation for 30 min at room temperature in the dark

Stop the reaction with 100 µl/well of R10

READING at 450/620 nm

The results can be interpreted by taking a THRESHOLD serum giving a response above which the sera tested would be considered as positive. This serum is chosen and diluted so as to give a significantly higher signal than the background noise.

2) DOUBLE EPITOPE ELISA test

a) Reagents

Preparation of the plates

The plates are sensitized with a solution of N protein at 1 µg/ml in a 10 mM PBS buffer, pH 7.2, phenol red at 0.25 ml/l. 100 µl of solution are deposited in the wells and left to incubate at room temperature overnight. Saturation is obtained by prewashing in 10 mM PBS/0.1% Tween buffer, followed by washing with a saturation solution 10 mM PBS, 25% (V/V) milk.

Diluent sera and conjugate
Buffer 50 mM TRIS saline, pH 8, 2% milk

Conjugate
This is the purified recombinant N protein coupled with peroxidase according to the Nakane protocol

(Nakane P.K. and Kawaoi A.; (1974): *Peroxydase-labeled antibody, a new method of conjugation. The Journal of Histochemistry and Cytochemistry* Vol. 22, N) 23, pp. 1084-1091), in respective molar ratios 1/2. This ProtN POD conjugate is used at a concentration of 2 µg/ml in serum/conjugate diluent.

Other solutions:
Washing solution R2, solutions for visualization with TMB R8, diluent, R9 chromogen, R10 stopping solution: reagents marketed by Bio-Rad (e.g. Platelia pylori kit, ref 72778).

b) Procedure

1st step in "predilution" plate

- Dilute each serum 1/5 in the predilution plate (48 µl of diluent + 12 µl of serum).
- After having diluted all the sera, distribute 60 µl of conjugate.
- Where appropriate, the serum + conjugate mix is left to incubate.

2nd step in "reaction" plate

- Transfer 100 µl of mixture/well into the reaction plate
- Incubation 1 h 37°C
- 5 washings in 10× WASHING solution R2 diluted 10-fold beforehand in demineralized water (→ 1× washing solution)
- Distribute 200 µl/well of visualization solution (to be diluted immediately before use e.g.: 1 ml of R9 in 10 ml of R8)
- Incubation 30 min at room temperature and protected from light
- Stop the reaction with 100 µl/well of R10
- READING at 450/620 nm

Likewise as for the indirect ELISA test, the results can be interpreted using a "threshold value" serum. Any serum having a response greater than the threshold value serum will be considered as positive.

2) Results

The sera of patients classified as probable cases of SARS from the French hospital of Hanoi, Vietnam or in relation with the French hospital of Hanoi (JYK) were analyzed using the indirect IgG-N test and the double epitope N test.

The results of the indirect IgG-N test (figures 14 and 15) and double epitope N test (figures 16 and 17) show an excellent correlation between them and with an indirect ELISA test comparing the reactivity of the sera toward a lysate of VeroE6 cells infected or not infected with SARS-CoV (ELISA-SARS-CoV lysate; see table V below). All the sera collected 12 days or more after the onset of the symptoms were found to be positive, including in patients for whom it had not been possible to document the SARS-CoV virus infection by analyzing respiratory samples by RT-PCR, probably because of a sample being collected too late during the infection (≥ D12). In the case of the patient TTH for whom a nasal sample collected on D7 was found to be negative by RT-PCR, the quality of the sample may be in question.

Some sera were found to be negative whereas the presence of SARS-CoV was detected by RT-PCR. They are in all cases early sera collected less than 10 days after the onset of the symptoms (e.g.: serum # 032637). In the case of a patient PTTH (serum # 032673), only a suspicion of SARS was raised at the time the samples were collected.

In conclusion, the indirect IgG-N and N-double epitope serological tests make it possible to document the SARS-CoV infection in all the patients for the sera collected 12 days or more after the infection.

Table V: Results of the ELISA tests

Sample Num	Patient	Day	PCR-SARS (1)	ELISA SARS-CoV lysate (2)	IgG-N (2nd series)	2Xepitope (2nd series)
033168	JYK	38	POS	+++	>5000	NT
033597	JYK	74	POS	NT	≈ 5000	NT
032552	VTT	8	NEG-D3&D8&D12	NEG	<200	<5
032544	CTP	16	NEG D16&D20	++	>5000	>>20
032546	CJF	15	NEG D15&D19	++	>5000	>>20
032548	PTL	17	NEG D17&D21	++	>5000	>>20
032550	NTH	17	NEG-D17&D21	++	>5000	>>20
032553	VTT	8	NEG-D3&D8&D12	NEG	<200	<5
032554	NTBV	4	POS	NEG	<200	<5
032555	NTBV	4	POS	NEG	<200	
032564	NTP	15	POS	++	>5000	>>20
032629	NVH	4	POS	NEG	<200	<5
032631	BTTX	9	POS	NEG	<200	<5
032635	NHH	4	POS	NEG	<200	<5
032637	NHB	10	POS	NEG	<200	<5
032642	BTTX	9	POS	NEG	<200	<5
032643	LTDH	1	POS	NEG	<200	<5
032644	NTBV	4	POS	NEG	<200	<5
032646	TTH	12	NEG D7&D12&D16	++	>5000	>>20
032647	DTH	17	NEG D17&D21	++	>5000	>>20
032648	NNT	15	NEG D15&D19	++	>5000	>>20
032649	PTH	17	NEG D17&D21	++	>5000	>>20
032672	LVV	16	NEG D16&D20	+	>5000	>>20
032673	PTTH	NA	NEG	NEG	<200	<5
032674	PNB	17	NEG D17&D21	++	>5000	>>20
032682	VTH	12	NEG D12&D16	++	>5000	>>20
032683	DTV	17	NEG D17&D21	+	>1000	>>20

Remarks:

(1): The RT-PCR analyses were carried out by nested

RT-PCR BNI, LC Artus and LC-N on nasal or pharyngeal swabs; POS means that at least one sample was found to be positive in this patient.

(2): The reactivity of the sera in the ELISA test using a lysate of cells infected with SARS-CoV was classified as very highly reactive (+++), highly reactive (++), reactive (+) and negative according to the OD value obtained at the dilutions tested.

Example 8: Detection of SARS-associated coronavirus (SARS-CoV) by RT-PCR

1) Real time development of RT-PCR conditions with the aid of primers specific for the gene for the nucleocapsid protein - "Light Cycler N" test

a) Design of the primers and probes

The primers and probes were designed from the sequence of the genome of the SARS-CoV strain derived from the sample recorded under the number 031589, with the aid of the programme "Light Cycler Probe Design (Roche)". Thus, the following two series of primers and probes were selected:

- **series 1 (SEQ ID NO: 60, 61, 64, 65):**
- sense primer: N/+/28507: 5'-GGC ATC GTA TGG GTT G-3' [28507-28522]
- antisense primer: N/-/28774: 5'-CAG TTT CAC CAC CTC C-3' [28774-28759]
- probe 1: 5'-GGC ACC CGC AAT CCT AAT AAC AAT GC-fluorescein 3' [28561-28586]
- probe 2: 5' Red705-GCC ACC GTG CTA CAA CTT CCT-phosphate [28588-28608]

- **series 2 (SEQ ID NO: 62, 63, 66, 67)**
- sense primer: N/+/28375: 5'-GGC TAC TAC CGA AGA G-3' [28375-28390]

- antisense primer: N/-/28702: 5'-AAT TAC CGC GAC TAC G-3' [28702-28687]
- probe 1: SARS/N/FL: 5'-ATA CAC CCA AAG ACC ACA TTG GC-fluorescein 3' [28541-28563]
- probe 2: SARS/N/LC705: 5' Red705-CCC GCA ATC CTA ATA ACA ATG CTG C-phosphate 3' [28565-28589]

b) Analysis of the efficacy of the two primer pairs

In order to test the respective efficacy of the two pairs of primers, an RT-PCR amplification was carried out on a synthetic RNA corresponding to nucleotides 28054-29430 of the genome of the SARS-CoV strain derived from the sample recorded under the number 031589 and containing the sequence of the N gene.

More specifically:

This synthetic RNA was prepared by *in vitro* transcription with the aid of the T7 phage RNA polymerase, of a DNA template obtained by linearization of the plasmid SRAS-N with the enzyme *Bam H1*. After eliminating the DNA template by digestion with the aid of DNAse 1, the synthetic RNAs are purified by a phenol-chloroform extraction, followed by two successive precipitations in ammonium acetate and isopropanol. They are then quantified by measuring the absorbance at 260 nm and their quality is checked by the ratio of the absorbances at 260 and 280 nm and by agarose gel electrophoresis. Thus, the concentration of the synthetic RNA preparation used for these studies is 1.6 mg/ml, which corresponds to 2.1×10^{15} copies/ml of RNA.

Decreasing quantities of synthetic RNA were amplified by RT-PCR with the aid of the "Superscript™ One-Step RT-PCR with Platinum® Taq" kit and the pairs of primers No. 1 (N/+/28507, N/-/28774) (figure 1A) and No. 2

(N/+/28375, N/-/28702) (figure 1B), according to the supplier's instructions. The amplification conditions used are the following: the cDNA was synthesized by incubation for 30 min at 45°C, 15 min at 55°C and then 2 min at 94°C and it was then amplified by 5 cycles comprising: a step of denaturation at 94°C for 15 sec, a step of annealing at 45°C for 30 sec and then a step of extension at 72°C for 30 sec, followed by 35 cycles comprising: a step of denaturation at 94°C for 15 sec, a step of annealing at 55°C for 30 sec and then a step of extension at 72°C for 30 sec, with 2 sec of additional extension at each cycle, and a final step of extension at 72°C for 5 min. The amplification products obtained were then kept at 10°C.

The results presented in figure 11 show that the pair of primers No. 2 (N/+/28375, N/-/28702) makes it possible to detect up to 10 copies of RNA (band of weak intensity) or 10^2 copies (band of good intensity) against 10^4 copies for the pair of primers No. 1 (N/+/28507, N/-/28774). The amplicons are respectively 268 bp (pair 1) and 328 bp (pair 2).

c) Development of real time RT-PCR

A real time RT-PCR was developed with the aid of the pair of primers No. 2 and of the pair of probes consisting of SRAS/N/FL and SRAS/N/LC705 (figure 2).

The amplification was carried out on a LightCycler™ (Roche) with the aid of the "Light Cycler RNA Amplification Kit Hybridization Probes" kit (reference 2 015 145, Roche) under the following optimized conditions. A reaction mixture containing: H_2O (6.8 µl), 25 mM $MgCl_2$ (0.8 µl, 4 µM Mg2+ final), 5X reaction mixture (4 µl), 3 µm probe SRAS/N/FL (0.5 µl, 0.075 µM final), 3 µM probe SRAS/N/LC705 (0.5 µl, 0.075 µM final), 10 µM primer N/+/28375 (1 µl, 0.5 µM

final), 10 µM primer N/-/28702 (1 µl, 0.5 µM final), enzyme mixture (0.4 µl) and sample (viral RNA, 5 µl) was amplified according to the following program:

- Reverse transcription: **50°C** 10:00 min analysis mode: none
- Denaturation: 95°C 30 sec × 1 analysis mode: none
- Amplification: 95°C 2 sec }
  50°C 15 sec analysis mode: quantification* } × 45
  72°C 13 sec thermal ramp 2.0°C/sec }
- Annealing: 40°C 30 sec × 1 analysis mode: none

* The fluorescence is measured at the end of the annealing and at each cycle (in SINGLE mode).

The results presented in figure 12 show that this real time RT-PCR is very sensitive since it makes it possible to detect 10^2 copies of synthetic RNA in 100% of the 5 samples analyzed (29/29 samples in 8 experiments) and up to 10 copies of RNA in 100% of the 5 samples analyzed (40/45 samples in 8 experiments). It also shows that this RT-PCR makes it possible to detect the presence of the SARS-CoV genome in a sample and to quantify the number of genomes present. By way of example, the viral RNA of a SARS-CoV stock cultured on Vero E6 cells was extracted with the aid of the "Qiamp viral RNA extraction" kit (Qiagen), diluted to 0.05×10^{-14} and analyzed by real time RT-PCR according to the protocol described above; the analysis presented in figure 12 shows that this virus stock contains 6.5×10^9 genome-equivalents/ml (geq/ml), which is entirely similar to the 1.0×10^{10} geq/ml value measured with the aid of the "RealArt™ HPA-Coronavirus LC RT PCR Reagents" kit marketed by Artus.

2) Development of nested RT-PCR conditions targeting the gene for RNA polymerase - "CDC (Centers for Disease

Control and Prevention)/IP nested RT-PCR" test

a) Extraction of the viral RNA

Clinical sample: QIAmp viral RNA Mini Kit (QIAGEN) according to the manufacturer's instructions, or an equivalent technique. The RNA is eluted in a volume of 60 µl.

b) "SNE/SAR" nested RT-PCR

First step: "SNE" coupled RT-PCR

The Invitrogen "*Superscript™ One-Step RT-PCR with Platinum® Taq*" kit was used, but the "*Titan*" kit from Roche Boehringer can be used in its place with similar results.

Oligonucleotides:
- SNE-S1 5' GGT TGG GAT TAT CCA AAA TGT GA 3'
- SNE-AS1 5' GCA TCA TCA GAA AGA ATC ATC ATG 3'

→ Expected size: **440 bp**

1. Prepare a mix:

H2O	6.5 µl
Reaction mix 2X	12.5 µl
Oligo SNE-S1 50 µM	0.2 µl
Oligo SNE-AS1 50 µM	0.2 µl
RNAsin 40 U/µl	0.12 µl
RT/Platinum Taq mix	0.5 µl

2. To 20 µl of the mix, add 5 µl of RNA and carry out the amplification on a thermocycler (ABI 9600 conditions):

2.1 45°C 30 min.
55°C 15 min.
94°C 2 min.
2.2. 94°C 15 sec. }

	45°C	30 sec.	} × 5 cycles
	72°C	30 sec.	}
2.3.	94°C	15 sec.	}
	55°C	30 sec.	} × 35 cycles
	72°C	30 sec. + 2 sec./cycle	}
2.4.	72°C	5 min.	
2.5	10°C	∞	

Storage at +4°C.

The RNAsin (N2511/N2515) from Promega was used as RNase inhibitors.

Synthetic RNAs served as positive control. As the control, 10^3, 10^2 and 10 copies of synthetic RNA R_{SNE} were amplified in each experiment.

Second step: "SAR" nested PCR

Oligonucleotides:
- SAR1-S 5' CCT CTC TTG TTC TTG CTC GCA 3'
- SAR1-AS 5' TAT AGT GAG CCG CCA CAC ATG 3'

→ Expected size: **121 bp**

1. Prepare a mix:

H2O	35.8 µl
Taq buffer 10X	5 µl
$MgCl_2$ 25 mM	4 µl
Mix dNTPs 5 mM	2 µl
Oligo SAR1-S 50 µM	0.5 µl
Oligo SAR1-AS 50 µM	0.5 µl
Taq DNA pol 5 U/µl	0.25 µl

AmpliTaq DNA Pol from Applied Biosystems was used (10X buffer without $MgCl_2$, ref 27216601).

2. To 48 µl of the mix, add 2 µl of the product from the first PCR and carry out the amplification (ABI 9600 conditions):

2.1.	94°C	2 min.	
2.2.	94°C	30 sec.	}
	45°C	45 sec.	} × 5 cycles
	72°C	30 sec.	}
2.3.	94°C	30 sec.	}
	55°C	30 sec.	} × 35 cycles
	72°C	30 sec. + 1 sec./cycle	}
2.4.	72°C	5 min.	
2.5	10°C	∞	

3. Analyze 10 µl of the reaction product on "*low-melting*" gel (Seakem GTG type) containing 3% agarose.

The sensitivity of the nested test is routinely, under the conditions described, 10 copies of RNA.

4. The fragments can then be purified on QIAquick PCR kit (QIAGEN) and sequenced with the oligos SAR1-S and SAR1-AS.

3) Detection of the SARS-CoV RNA by PCR from respiratory samples

a) First comparative study

A comparative study was carried out on a series of respiratory samples received by the National Reference Center for the Influenza Virus (Northern region) and likely to contain SARS-CoV. To do this, the RNA was extracted from the samples with the aid of the "Qiamp viral RNA extraction" kit (Qiagen) and analyzed by real time RT-PCR, on the one hand with the aid of the pairs of primers and probes of the No. 2 series under the conditions described above on the one hand, and on the other hand with the aid of the kit "LightCycler SARS-CoV quantification kit" marketed by Roche (reference 03 604 438). The results are summarized in table VI below. They show that 18 of the 26 samples are negative

and 5 of the 26 samples are positive for the two kits, while one sample is positive for the Roche kit alone and two for the "series 2" N reagents alone. Additionally, for 3 samples (20032701, 20032712, 20032714) the quantities of RNA detected are markedly higher with the reagents (probes and primers) of the No. 2 series. These results indicate that the "series 2" N primers and probes are more sensitive for the detection of the SARS-CoV genome in biological samples than those of the kit currently available.

Table VI: Real time RT-PCR analysis of the RNAs extracted from a series of samples from 5 patients with the aid of the pairs of primers and probes of the No. 2 series ("series 2" N) or of the kit "Lightcycler SARS-CoV quantification kit" (Roche). The type of sample is indicated as well as the number of copies of viral genome measured in each of the two tests. NEG: negative RT-PCR.

Sample No.	Patient	Type of sample	ROCHE KIT	"Series 2" N
20033082	K	nasal	NEG	NEG
20033083	K	pharyngeal	NEG	NEG
20033086	K	nasal	NEG	NEG
20033087	K	pharyngeal	NEG	NEG
20032802	M	nasal	NEG	NEG
20032803	M	expectoration	NEG	NEG
20032806	M	nasal or pharyngeal	NEG	NEG
20031746ARN2	C	pharyngeal	NEG	NEG
20032711	C	nasal or pharyngeal	39	NEG
20032910	B	nasal	NEG	NEG
20032911	B	pharyngeal	NEG	NEG
20033356	V	expectoration	NEG	NEG
20033357	V	expectoration	NEG	NEG
20031725	K	endotracheal asp.	NEG	150
20032657	K	endotracheal asp.	NEG	NEG
20032698	K	endotracheal asp.	NEG	NEG
20032720	K	endotracheal asp.	3	5
20033074	K	stools	115	257
20032701	M	pharyngeal	443	1676
20032702	M	expectoration	NEG	249
20031747ARN2	C	pharyngeal	NEG	NEG
20032712	C	unknown	634	6914
20032714	C	pharyngeal	17	223
20032800	B	nasal	NEG	NEG
20033353	V	nasal	NEG	NEG
20033384	V	nasal	NEG	NEG

b) Second comparative study

The performance of various nested RT-PCR and real time RT-PCR methods were then compared for 121 respiratory samples from possible cases of SARS at the French hospital in Hanoi, Vietnam, taken between the 4th and the 17th day after the onset of the symptoms. Among these samples, 14 were found to be positive during a first test using the nested RT-PCR method targeting ORF1b (encoding replicase) as described initially by Bernhard Nocht Institute (BNI nested RT-PCR). Information relating to this test is available on the internet, at the address http://www15.bni-hamburg.de/bni2/neu2/getfile.acgi?area_engl=diagnostics&pid=4112.

The various tests compared in this study are:

- the quantitative RT-PCR method according to the invention, with the "series 2" N primers and probes described above (LightCycler N column),
- the nested RT-PCR test targeting the RNA polymerase gene described above, developed by the CDC, BNI and Institut Pasteur (CDC/IP nested RT-PCR),
- the ARTUS kit with the reference "HPA Corona LC RT-PCR Kit # 5601-02", which is a real time RT-PCR test targeting the ORF1b gene,
- the BNI nested RT-PCR test, also targeting the RNA polymerase gene mentioned above.

The inventors observed:

1) an inter-test variability for the same technique, linked to the degradation of the RNA preparation during repeated thawing, in particular for the samples containing the lowest quantities of RNA,

2) a reduced sensitivity of the CDC/IP nested RT-PCR

compared with the BNI nested RT-PCR, and

3) a comparable sensitivity of the quantitative RT-PCR test according to the invention (LightCycler N) compared with the Artus LightCycler (LC) test.

These results, which are presented in table VII below, show that the quantitative RT-PCR test according to the invention constitutes an excellent addition - or an alternative - to the tests currently available. Indeed, the SARS-linked coronavirus is an emergent virus which is capable of changing rapidly. In particular, the gene for the RNA polymerase of the SARS-linked coronavirus, which is targeted in most of the tests currently available, can recombine with that of other coronaviruses not linked to SARS. The use of a test targeting this gene exclusively could then lead to the production of false-negatives.

The quantitative RT-PCR test according to the invention does not target the same genomic region as the ARTUS kit since it targets the gene encoding the N protein. By carrying out a diagnostic test targeting two different genes of the SARS-linked coronavirus, it can therefore be hoped to avoid false-negative type results which could be due to the genetic evolution of the virus.

Furthermore, it appears particularly advantageous to target the gene for the nucleocapsid protein because it is very stable because of the high selection pressure linked to the high structural constraints regarding this protein.

<u>Table VII</u>: Comparison of various methods of analysis by gene amplification, from 121 samples of probable cases of SARS at the French hospital in Hanoi, Vietnam (epidemic 2003)

NRC No.	Sample type (1)	Sample collection day	Patient	CDC/IP nested RT-PCR	BNI nested RT-PCR	Artus Light Cycler kit	Light Cycler N (IP)
107 samples	N and P			Negative	Negative	Negative	Negative
032529	P	10	NHB	Negative	Positive	Negative	Negative
032530	N	10	NHB	Positive	Positive	3.10E+01	4.20E+01
032531	P	7	LP	Positive	Positive	7.70E+00	3.10E+00
032534	N	15	BND	Positive	Positive	1.60E+00	Negative
032600	P	4	NHH	Negative	Positive	Negative	1.30E+02
032612	P	17	NTS	Negative	Positive	Negative	Negative
032688	P	9	BTX	Positive	Positive	Negative	Negative
032689	N	4	NVH	Positive	Positive	1.20E+01	2.30E+02
032690	P	4	NVH	Negative	Positive	1.60E+00	Negative
032727	P	8	NVH	Positive	Positive	2.30E+02	4.00E+02
032728	N	8	NVH	Positive	Positive	1.10E+03	1.60E+04
032729	P	14	NHB	Positive	Positive	5.80E+00	3.40E+01
032730	N	14	NHB	Positive	Positive	1.30E+02	4.80E+02
032741	P	8	NHH	Positive	Positive	2.10E+02	1.30E+02
			positives	10	14	10	9
			fraction detected from the 14 positives	71.4%	100.0%	71.4%	64.3%

(1) P = pharyngeal swab
N = nasal swab

Example 9: Production and characterization of monoclonal antibodies directed against the N protein

Balb C mice were immunized with the purified recombinant N protein and their spleen cells fused with an appropriate murine myeloma according to the Köhler and Milstein techniques.

Nineteen anti-N antibody secreting hybridomas were preselected and their immunoreactivities determined. These antibodies do indeed recognize the recombinant N protein (in ELISA) with variable intensities, and the natural viral N protein in ELISA and/or in Western blotting. Figures 18 to 20 show the results of these tests for 15 of these 19 monoclonal antibodies.

The highly reactive clones 12, 17, 28, 57, 72, 76, 86, 87, 98, 103, 146, 156, 166, 170, 199, 212, 218, 219 and 222 were subcloned. Specificity studies were carried out with the appropriate tools in order to determine

the epitopes recognized and verify the absence of reactivity toward other human coronaviruses and certain respiratory viruses.

Epitope mapping studies (performed on spot membrane with the aid of overlapping peptides of 15 aa) and additional studies performed on the natural N protein in Western blotting revealed the existence of 4 groups of monoclonal antibodies:

1. Monoclonal antibodies specific for a major linear epitope at the N-ter position (75-81, sequence: INTNSVP).

The representative of this group is antibody 156. The hybridoma producing this antibody was deposited at the Collection Nationale de Cultures de Microorganismes (CNCM) of the Institut Pasteur (Paris, France) on December 1, 2004, under the number I-3331. This same epitope is also recognized by a rabbit serum (anti-N polyclonal) obtained by conventional immunization with the aid of this same N protein.

2. Monoclonal antibodies specific for a major linear epitope located in a central position (position 217-224, sequence: ETALALL); the representatives of this group are the monoclonal antibodies 87 and 166. The hybridoma producing antibody 87 was deposited at the CNCM on December 1, 2004, under the number I-3328.

3. Monoclonal antibodies specific for a major linear epitope located at the C-terminal position (position 403-408, sequence: DFFRQL), the representatives of this group are the antibodies 28, 57 and 143. The hybridoma producing antibody 57 was deposited at the CNCM on December 1, 2004, under the number I-3330.

4. Monoclonal antibodies specific for a discontinuous

conformational epitope. This group of antibodies does not recognize any of the peptides spanning the sequence of the N protein, but react strongly on the non-denatured natural protein. The representative of this final group is the antibody 86. The hybridoma producing this antibody was deposited at the CNCM on December 1, 2004, under the number I-3329.

Table VIII below summarizes the epitope mapping results obtained:

Table VIII: Epitope mapping of the monoclonal antibodies

Antibody	Epitope	Position	Region
28	DFSRQL *Q*	403 ... 408	C – Ter.
143	DFSRQL *Q*		
76	DFSRQL *Q*		
57	DFSRQL *Q*		
	FFGMS RI	315 ... 319	
146	LPQRQ	383 ...387	
166	ETALALL*LL*	217 ... 224	central
87	ETALALL	217... 224	
156	*INTNSGP*	*75... 81*	*N – Ter.*
86	*Conformational*		
212	*Conformational*		
170	Conformational		

In addition, as illustrated in particular in figures 18 and 19, these antibodies exhibit no reactivity in ELISA and/or in WB toward the N protein of the human corona-virus 229 E.

Example 10: Combinations of the monoclonal antibodies for the development of a sensitive immunocapture test specific for the viral N antigen in the serum or

biological fluids of patients infected with the SARS-CoV virus

The antibodies listed below were selected because of their very specific properties for an additional capture and detection study of the viral N protein, in the serum of the subjects or patients.

These antibodies were produced in ascites on mice, purified by affinity chromatography and used alone or in combination, as capture antibodies and as signal antibodies.

List of the antibodies selected:

- Ab anti-C-ter region (No. 28, 57, 143)
- Ab anti-central region (No. 87, 166)
- Ab anti-N-ter region (No. 156)
- Ab anti-discontinuous conformational epitope (86)

1) Preparation of the reagents:

a) Immunocapture ELISA plates

The plates are sensitized with the antibody solutions at 5 µg/ml in 0.1 M carbonate buffer, pH 9.6. The (monovalent or plurivalent) solutions are deposited in a volume of 100 µl in the wells and incubated overnight at room temperature. These plates are then washed with PBS buffer (10 mM pH 7.4 supplemented with 0.1% Tween 20) and then saturated with a PBS solution supplemented with 0.3% BSA and 5% sucrose). The plates are then dried and then packaged in a bag in the presence of a desiccant. They are ready to use.

b) Conjugates

The purified antibodies were coupled with peroxidase according to the Nakane protocol (Nakane et al. - 1974,

J. of Histo and cytochemistry, vol. 22, pp. 1084-1091) in a ratio of one molecule of IgG per 3 molecules of peroxidase. These conjugates were purified by exclusion chromatography and stored concentrated (concentration between 1 and 2 mg/ml) in the presence of 50% glycerol and at -20°C. They are diluted for their use in the assays at the final concentration of 1 or 2 µg/ml in PBS buffer (pH 7.4) supplemented with 1% BSA.

c) Other reagents

- Human sera negative for all the serum markers for the HIV, HBV, HCV and THLV viruses
- Pool of negative human sera supplemented with 0.5% Triton X 100
- Inactivated viral Ag: viral culture supernatant inactivated by irradiation and inactivation verified after placing in culture on sensitive cells - titer of the suspension before inactivation about 10^7 infectious particles per ml or alternatively about 5×10^9 physical viral particles per ml of antigen
- The Ag samples diluted in negative human serum: these samples were prepared by diluting 1:100 and then by 5-fold serial dilution.
  These noninfectious samples mimic human samples thought to contain low to very low concentrations of viral nucleoprotein N. Such samples are not available for routine work.
- Washing solution R2, solution for visualization TMB R8, chromogen R9 and stop solution R10, are the generic reagents marketed by Bio-Rad in its ELISA kits (e.g.: *Platelia pylori* kit ref. 72778).

2) Procedure

The samples of human sera overloaded with inactivated viral Ag are distributed in an amount of 100 µl per

well, directly in the ready-to-use sensitized plates, and then incubated for 1 hour at 37°C (Bio-Rad IPS incubation).

The material not bound to the solid phase is removed by 3 washings (washing with dilute R2 solution, automatic LP 35 washer).

The appropriate conjugates, diluted to the final concentration of 1 or 2 µg/ml, are distributed in an amount of 100 µl per well and the plates are again incubated for one hour at 37°C (IPS incubation).

The excess conjugate is removed by 4 successive washings (dilute R2 solution - LP 35 washer).

The presence of conjugate attached to the plates is visualized after adding 100 µl of visualization solution prepared before use (1 ml of R9 and 10 ml of R8) and after incubation for 30 minutes, at room temperature and protected from light.

The enzymatic reaction is finally blocked by adding 100 µl of R10 reagent (1 N H_2SO_4) to all the wells.

The reading is carried out with the aid of an appropriate microplate reader at double wavelength (450/620 nm).

The results can be interpreted by using, as provisional threshold value, the mean of at least two negative controls multiplied by a factor of 2 or alternatively the mean of 100 negative sera supplemented with an increment corresponding to 6 SD (*standard deviation* calculated on the 100 individual measurements).

3) Results

Various capture antibody and signal antibody combinations were tested based on the properties of the antibodies selected, and avoiding the combinations of antibodies specific for the same epitopes in solid phase and as conjugates.

The best results were obtained with the 4 combinations listed below. These results are reproduced in table IX below.

1. Combination F/28
Solid phase (Ab 166 + 87 central region): conjugate antibody 28 (C-ter)

2. Combination G/28
Solid phase (Ab 86 - conformational epitope): conjugate antibody 28 (C-ter)

3. Combination H/28
Solid phase (Ab 86, 166 and 87 central region and conformational epitope): conjugate antibody 28 (C-ter)

4. Combination H/28 + 87
Solid phase (Ab 86, 166 and 87 central region and conformational epitope): mixed conjugate antibodies 28 (C-ter) and 87 (central)

5. Combination G/87
Solid phase (Ab 86 - conformational epitope): conjugate antibody 87 (central region)

The first 4 combinations exhibit equivalent and reproduced performance levels, greater than the other combinations used (such as for example the combination G/87). Of course, in these combinations, a monoclonal antibody may be replaced with another antibody recognizing the same epitope. Thus, the following variants may be mentioned:

6. Variant of the combination F/28
Solid phase (Ab 87 only): conjugate antibody 57 (C-ter)

7. Variant of the combination G/28
Solid phase (Ab 86 - conformational epitope): conjugate antibody 57 (C-ter)

8. Variant of the combination H/28
Solid phase (Ab 86 and 87 central region and conformational epitope): conjugate antibody 57 (C-ter)

9. Variant of the combination H/28 + 87
Solid phase (Ab 86 and 87 central region and conformational epitope): mixed conjugate antibodies 57 (C-ter) and 87 (central)

Table IX: Test of immunoreactivity of the anti-SARS-CoV nucleoprotein Abs: optical densities measured with each combination of antibodies according to the dilutions of the inactivated viral antigen.

No.	Dilution	F/28	G/28	G/87	H/28	H/28+87
0	1/100	5	5	3.495	3.900	5
1	1/500	3.795	3.814	1.379	3.702	3.804
2	½ 500	2.815	2.950	0.275	3.268	2.680
3	1/12 500	0.987	1.038	0.135	1.374	0.865
4	1/62 500	0.404	0.348	0.125	0.480	0.328
5	1/312 500	0.285	0.211	0.123	0.240	0.215
6	Control	0.210	0.200	0.098	0.186	0.156
7	Control	0.269	0.153	0.104	0.193	0.202

The detection limit for these 4 experimental trials corresponds to the antigen dilution in negative serum 1:62 500. A rapid extrapolation suggests the detection of less than 10^3 infectious particles per ml of sera.

From this study, it is evident that the most appropriate antibodies for the capture of the native viral nucleoprotein are the antibodies specific for the

central region and/or for a conformational epitope, both being antibodies also selected for their high affinity for the native antigen.

Having determined the best antibodies for the composition of the solid phase, the antibodies to be selected as a priority for the detection of the antigens attached to the solid phase are the complementary antibodies specific for a dominant epitope in the C-ter region. The use of any other complementary antibody specific for epitopes located in the N-ter region of the protein leads to average or poor results.

Example 11: Eukaryotic expression systems for the SARS-associated coronavirus (SARS-CoV) spicule (S) protein

1) Optimization of the conditions for expression of the SARS-CoV S in mammalian cells

The conditions for transient expression of the SARS-CoV spicule (S) protein were optimized in mammalian cells (293T, VeroE6).

For that, a DNA fragment containing the cDNA for SARS-CoV S was amplified by PCR with the aid of the oligonucleotides 5'-ATAGGATCCA CCATGTTTAT TTTCTTATTA TTTCTTACTC TCACT-3' and 5'-ATACTCGAGTT ATGTGTAATG TAATTTGACA CCCTTG-3' from the plasmid pSARS-S (C.N.C.M. No. I-3059) and then inserted between the BamH1 and Xho1 sites of the plasmid pTRIPΔU3-CMV containing a lentiviral vector TRIP (Sirven, 2001, Mol. Ther., 3, 438-448) in order to obtain the plasmid pTRIP-S. The BamH1 and Xho1 fragment containing the cDNA for S was then subcloned between BamH1 and Xho1 of the eukaryotic expression plasmid pcDNA3.1(+) (Clontech) in order to obtain the plasmid pcDNA-S. The Nhe1 and Xho1 fragment containing the cDNA for S was then subcloned between

the corresponding sites of the expression plasmid pCI (Promega) in order to obtain the plasmid pCI-S. The WPRE sequences of the woodchuck hepatitis virus ("*Woodchuck Hepatitis Virus posttranscriptional regulatory element*") and the CTE sequences ("*constitutive transport element*") of the simian retrovirus from Mason-Pfizer were inserted into each of the two plasmids pcDNA-S and pCI-S between the Xho1 and Xba1 sites in order to obtain respectively the plasmids pcDNA-S-CTE, pcDNA-S-WPRE, pCI-S-CTE and pCI-S-WPRE (figure 21). The plasmid pCI-S-WPRE was deposited at the CNCM, on November 22, 2004, under the number I-3323. All the inserts were sequenced with the aid of a BigDye Terminator v1.1 kit (Applied Biosystems) and an automated sequencer ABI377.

The capacity of the plasmid constructs to direct the expression of SARS-CoV S in mammalian cells was assessed after transfection of VeroE6 cells (figure 22). In this experiment, monolayers of 5×10^5 VeroE6 cells in 35 mm Petri dishes were transfected with 2 μg of plasmids pcDNA (as control), pcDNA-S, pCI and pCI-S and 6 μl of Fugene6 reagent according to the manufacturer's instructions (Roche). After 48 hours of incubation at 37°C and under 5% CO_2, cellular extracts were prepared in loading buffer according to Laemmli, separated on 8% SDS polyacrylamide gel, and then transferred onto a PVDF membrane (BioRad). The detection of this immunoblot (Western blot) was carried out with the aid of an anti-S rabbit polyclonal serum (immune serum from the rabbit P11135: cf. example 4 above) and donkey polyclonal antibodies directed against rabbit IgGs and coupled with peroxidase (NA934V, Amersham). The bound antibodies were visualized by luminescence with the aid of the ECL+ kit (Amersham) and autoradiography films Hyperfilm MP (Amersham).

This experiment (figure 22) shows that the plasmid pcDNA-S does not make it possible to direct the expression of SARS-CoV S at detectable levels whereas the plasmid pCI-S allows a weak expression, close to the limit of detection, which may be detected when the film is overexposed. Similar results were obtained when the expression of S was sought by immunofluorescence (data not shown). This impossibility to detect effective expression of S cannot be attributed to the detection techniques used since the S protein can be detected at the expected size (180 kDa) in an extract of cells infected with SARS-CoV or in an extract of VeroE6 cells infected with the recombinant vaccinia virus VV-TF7.3 and transfected with the plasmid pcDNA-S. In this latter experiment, the virus VV-TF7.3 expresses the RNA polymerase of the T7 phage and allows the cytoplasmic transcription of an uncapped RNA capable of being efficiently translated. This experiment suggests that the expression defects described above are due to an intrinsic inability of the cDNA for S to be efficiently expressed when the step for transcription to messenger RNA is carried out at the nuclear level.

In a second experiment, the effect of the CTE and WPRE signals on the expression of S was assessed after transfection of VeroE6 (figure 23A) and 293T (figure 23B) cells and according to a protocol similar to that described above. Whereas the expression of S cannot be detected after transfection of the plasmids pcDNA-S-CTE and pcDNA-S-WPRE derived from pcDNA-S, the insertion of the WPRE and CTE signals greatly improves the expression of S in the context of the expression plasmid pCI-S.

To specify this result, a second series of experiments were carried out where the immunoblot is quantitatively visualized by luminescence and acquisition on a digital

imaging device (FluorS, BioRad). The analysis of the results obtained with the QuantityOne v4.2.3 software (BioRad) shows that the WPRE and CTE sequences increase respectively the expression of S by a factor of 20 to 42 and 10 to 26 in Vero E6 cells (table X). In 293T cells (table X), the effect of the CTE sequence is more moderate (4 to 5 times) whereas that of the WPRE sequence remains high (13 to 28 times).

Table X Quantitative analysis of the effect of the CTE and WPRE signals on the expression of SARS-CoV S: Cellular extracts were prepared 48 hours after transfection of VeroE6 or 293T cells with the plasmid pCI, pCI-S, pCI-S-CTE and pCI-S-WPRE and analyzed by Western blotting as described in the legend to figure 22. The Western blot is visualized by luminescence (ECL+, Amersham) and acquisition on a digital imaging device (FluorS, BioRad). The expression levels are indicated according to an arbitrary scale where the value of 1 represents the level measured after transfection of the plasmid pCI-S.
Two independent experiments were carried out for each of the two cell types. In experiment 1 on VeroE6 cells, the transfections were carried out in duplicate and the results are indicated in the form of the mean and standard deviation values for the expression levels measured.

Plasmid	cell	exp. 1	exp. 2
PCI	VeroE6	0.0	0.0
pCI-S	VeroE6	1.0±0.1	1.0
pCI-S-CTE	VeroE6	9.8 ±0.9	26.4
pCI-S-WPRE	VeroE6	20.1±2.0	42.3
PCI	293T	0.0	0.0
PCI-S	293T	1.0	1.0
PCI-S-CTE	293T	4.6	4.0
PCI-S-WPRE	293T	27.6	12.8

In summary, all these results show that the expression, in mammalian cells, of the cDNA for the SARS-CoV S under the control of the RNA polymerase II promoter

sequences requires, to be efficient, the expression of a splice signal and of either of the sequences WPRE and CTE.

2) Production of stable lines allowing the expression of SARS-CoV S

The cDNA for the SARS-CoV S protein was cloned in the form of a BamH1-Xho1 fragment into the plasmid pTRIPΔU3-CMV containing a defective lentiviral vector TRIP with central DNA flap (Sirven et al., 2001, Mol. Ther., 3: 438-448) in order to obtain the plasmid pTRIP-S (figure 24). Transient cotransfection according to Zennou et al. (2000, Cell, 101: 173-185) of this plasmid, of an encapsidation plasmid (p8.2) and of a plasmid for expression of the VSV envelope glycoprotein G (pHCMV-G) in 293T cells allowed the preparation of retroviral pseudoparticles containing the vector TRIP-S and pseudotyped with the envelope protein G. These pseudotyped TRIP-S vectors were used to translate 293T and FRhK-4 cells: no expression of the S protein could be detected by Western blotting and immunofluorescence in the transduced cells (data not presented).

The optimum expression cassettes consisting of the CMV virus immediate/early promoter, a splice signal, cDNA for S and either of the posttranscriptional signals WPRE or CTE described above were then substituted for the EF1α-EGFP cassette of the defective lentiviral expression vector with central DNA flap TRIPΔU3-EF1α (Sirven et al., 2001, Mol. Ther., 3: 438-448) (figure 25). These substitutions were carried out by a series of successive subclonings of the S expression cassettes which were excised from the plasmids pCT-S-CTE (BglII-Apa1) or respectively pCI-S-WPRE (BglII-Sal1) and then inserted between the Mlu1 and Kpn1 sites or respectively Mlu1 or Xho1 sites of the plasmid TRIPΔU3-EF1α in order to obtain the plasmids pTRIP-SD/SA-S-CTE and pTRIP-SD/SA-S-WPRE, deposited at the

CNCM, on December 1, 2004, under the numbers I-3336 and I-3334, respectively. Pseudotyped vectors were produced according to Zennou et al. (2000, Cell, 101: 173-185) and used to transduce 293T cells (10 000 cells) and FRhK-4 cells (15 000 cells) according to a series of 5 successive transduction cycles with a quantity of vectors corresponding to 25 ng (TRIP-SD/SA-S-CTE) or 22 ng TRIP-SD/SA-S-WPRE) of p24 per cycle.

The transduced cells were cloned by limiting dilution and a series of clones were qualitatively analyzed for the expression of SARS-CoV S by immunofluorescence (data not shown), and then quantitatively by Western blotting (figure 25) with the aid of an anti-S rabbit polyclonal serum. The results presented in figure 25 show that clones 2 and 15 of FrhK4-s-CTE cells transduced with TRIP-SD/SA-S-CTE and clones 4, 9 and 12 of FRhK4-S-WPRE cells transduced with TRIP-SD/SA-S-WPRE allow the expression of the SARS-CoV S at respectively low or moderate levels if they are compared to those which can be observed during infection with SARS-CoV.

In summary, the vectors TRIP-SD/SA-S-CTE and TRIP-SD/SA-S-WPRE allow the production of stable clones of FRhK-4 cells and similarly 293T cells expressing SARS-CoV S, whereas the assays carried out with the "parent" vector TRIP-S remained unsuccessful, which demonstrates the need for a splice signal and for either of the sequences CTE and WPRE for the production of stable cell clones expressing the S protein.

In addition, these modifications of the vector TRIP (insertion of a splice signal and of a post-transcriptional signal like CTE and WPRE) could prove advantageous for improving the expression of other cDNAs than that for S.

3) Production of stable lines allowing the expression

of a soluble form of SARS-CoV S. Purification of this recombinant antigen.

A cDNA encoding a soluble form of the S protein (Ssol) was obtained by fusing the sequences encoding the ectodomain of the protein (amino acids 1 to 1193) with those of a tag (FLAG:DYKDDDDK) via a BspE1 linker encoding the SG dipeptide. Practically, in order to obtain the plasmid pcDNA-Ssol, a DNA fragment encoding the ectodomain of SARS-CoV S was amplified by PCR with the aid of the oligonucleotides 5'-ATAGGATCCA CCATGTTTAT TTTCTTATTA TTTCTTACTC TCACT-3' and 5'-ACCTCCGGAT TTAATATATT GCTCATATTT TCCCAA-3' from the plasmid pcDNA-S, and then inserted between the unique BamH1 and BspE1 sites of a modified eukaryotic expression plasmid pcDNA3.1(+) (Clontech) containing the tag sequence FLAG between its BamH1 and Xho1 sites:

```
// GGATCC ...nnn... TCC GGA GAT TAT AAA GAT GAC GAC GAT AAA TAA
    BamH1            S   G   D   Y   K   D   D   D   D   K  ter
CTCGAG //
Xho1
```

The Nhe1-Xho1 and BamH1-Xho1 fragments, containing the cDNA for S, were then excised from the plasmid pcDNA-Ssol, and subcloned between the corresponding sites of the plasmid pTRIP-SD/SA-S-CTE and of the plasmid pTRIP-SD-SA-S-WPRE, respectively, in order to obtain the plasmids pTRIP-SD/SA-Ssol-CTE and pTRIP-SD/SA-Ssol-WPRE, deposited at the CNCM, on December 1, 2004, under the numbers I-3337 and I-3335, respectively.

Pseudotyped vectors were produced according to Zennou et al. (2000, Cell, 101:173-185) and used to transduce FRhK-4 cells (15 000 cells) according to a series of 5 successive transduction cycles (15 000 cells) with a quantity of vector corresponding to 24 ng (TRIP-SD/SA-Ssol-CTE) or 40 ng (TRIP-SD/SA-Ssol-WPRE) of p24 per

cycle. The transduced cells were cloned by limiting dilution and a series of 16 clones transduced with TRIP-SD/SA-Ssol-CTE and of 15 clones with TRIP-SD/SA-Ssol-WPRE were analyzed for the expression of the Ssol polypeptide by Western blotting visualized with an anti-FLAG monoclonal antibody (figure 26 and data not presented), and by capture ELISA specific for the Ssol polypeptide which was developed for this purpose (table XI and data not presented). Part of the process for selecting the best secretory clones is shown in figure 26. Capture ELISA is based on the use of solid phases coated with polyclonal antibodies of rabbits immunized with purified and inactivated SARS-CoV. These solid phases allow the capture of the Ssol polypeptide secreted into the cellular supernatants, whose presence is then visualized with a series of steps successively involving the attachment of an anti-FLAG monoclonal antibody (M2, SIGMA), of anti-mouse IgG(H+L) biotinylated rabbit polyclonal antibodies (Jackson) and of a streptavidin-peroxidase conjugate (Amersham) and then the addition of chromogen and substrate (TMB + H_2O_2, KPL).

Table XI: Analysis of the expression of the Ssol polypeptide by cell lines transduced with the lentiviral vectors TRIP-SD/SA-Ssol-WPRE and TRIP-SD/SA-Ssol-CTE. The secretion of the Ssol polypeptide was assessed in the supernatant of a series of cell clones isolated after transduction of FRhK-4 cells with the lentiviral vectors TRIP-SD/SA-Ssol-WPRE and TRIP-SD/SA-Ssol-CTE. The supernatants diluted 1/50 were analyzed by a capture ELISA test specific for SARS-CoV S.

Vector	Clone	OD (450 nm)
Control	-	0.031
TRIP-SD/SA-Ssol-CTE	CTE2	0.547
	CTE3	0.668
	CTE9	0.171

	CTE12	0.208
	CTE13	0.133
TRIP-SD/SA-Ssol-WPRE	WPRE1	0.061
	WPRE10	0.134

The cell line secreting the highest quantities of Ssol polypeptide in the culture supernatant is the FRhK4-Ssol-CTE3 line. It was subjected to a second series of 5 cycles of transduction with the vector TRIP-SD/SA-Ssol-CTE under conditions similar to those described above and then cloned. The subclone secreting the highest quantities of Ssol was selected by a combination of Western blot and capture ELISA analysis: it is the subclone FRhK4-Ssol-30, which was deposited at the CNCM, on November 22, 2004, under the name I-3325.

The FRhK4-Ssol-30 line allows the quantitative production and purification of the recombinant Ssol polypeptide. In a typical experiment where the experimental conditions for growth, production and purification were optimized, the cells of the FRhK4-Ssol-30 line are inoculated in standard culture medium (pyruvate-free DMEM containing 4.5 g/l of glucose and supplemented with 5% FCS, 100 U/ml of penicillin and 100 µg/ml of streptomycin) in the form of a subconfluent monolayer (1 million cells per each 100 cm^2 in 20 ml of medium). At confluence, the standard medium is replaced with the secretion medium where the quantity of FCS is reduced to 0.5% and the quantity of medium reduced to 16 ml per each 100 cm^2. The culture supernatant is removed after 4 to 5 days of incubation at 35°C and under 5% CO_2. The recombinant polypeptide Ssol is purified from the supernatant by the succession of steps of filtration on 0.1 µm polyethersulfone (PES) membrane, concentration by ultrafiltration on a PES membrane with a 50 kD cut-off, affinity chromatography on anti-FLAG matrix with

elution with a solution of FLAG peptide (DYKDDDDK) at 100 µg/ml in TBS (50 mM tris, pH 7.4, 150 mM NaCl) and then gel filtration chromatography in TBS on sephadex G-75 beads (Pharmacia). The concentration of the purified recombinant Ssol polypeptide was determined by micro-BCA test (Pierce) and then its biochemical characteristics analyzed.

Analysis by 8% SDS acrylamide gel stained with silver nitrate demonstrates a predominant polypeptide whose molecular mass is about 180 kD and whose degree of purity may be evaluated at 98% (figure 27A). Two main peaks are detected by SELDI-TOF mass spectrometry (Cyphergen): they correspond to single and double charged forms of a predominant polypeptide whose molecular mass is thus determined at 182.6 ± 3.7 kD (figures 27B and C). After transfer onto Prosorb membrane and rinsing in 0.1% TFA, the N-terminal end of the Ssol polypeptide was sequenced in liquid phase by Edman degradation on 5 residues (ABI494, Applied Biosystems) and determined as being SDLDR (figure 27D). This demonstrates that the signal peptide located at the N-terminal end of the SARS-CoV S protein, composed of aa 1 to 13 (MFIFLLFLTLTSG) according to an analysis carried out with the software signalP v2.0 (Nielsen et al., 1997, *Protein Engineering,* 10:1-6), is cleaved from the mature Ssol polypeptide. The recombinant Ssol polypeptide therefore consists of amino acids 14 to 1193 of the SARS-CoV S protein fused at the C-terminals with a sequence SGDYKDDDDK containing the sequence of the FLAG tag (underlined). The difference between the theoretical molar mass of the naked Ssol polypeptide (132.0 kD) and the real molar mass of the mature polypeptide (182.6 kD) suggests that the Ssol polypeptide is glycosylated.

A preparation of purified Ssol polypeptide, whose protein concentration was determined by micro-BCA test,

makes it possible to prepare a calibration series in order to measure, with the aid of the capture ELISA test described above, the concentrations of Ssol present in the culture supernatants and to review the characteristics of the secretory lines. According to this test, the FRhK4-Ssol-CT3 line secretes 4 to 6 μg/ml of polypeptide Ssol while the FRhK4-Ssol-30 line secretes 9 to 13 μg/ml of Ssol after 4 to 5 days of culture at confluence. In addition, the purification scheme presented above makes it possible routinely to purify from 1 to 2 mg of Ssol polypeptide per liter of culture supernatant.

Example 12: Gene immunization involving the SARS-associated corona virus (SARS-CoV) spicule (S) protein

The effect of a splice signal and of the posttranscriptional signals WPRE and CTE was analyzed after gene immunization of BALB/c mice (figure 28).

For that, BALB/c mice were immunized at intervals of 4 weeks by injecting into the *tibialis anterior* a saline solution of 50 μg of plasmid DNA of pcDNA-S and pCI-S and, as a control, 50 μg of plasmid DNA of pcDNA-N (directing the expression of SARS-CoV N) or of pCI-HA (directing the expression of the HA of the influenza virus A/PR/8/34) and the immune sera collected 3 weeks after the 2nd injection. The presence of antibodies directed against the SARS-CoV S was assessed by indirect ELISA using as antigen a lysate of VeroE6 cells infected with SARS-CoV and, as a control, a lysate of noninfected VeroE6 cells. The anti-SARS-CoV antibody titers (TI) are calculated as the reciprocal of the dilution producing a specific OD of 0.5 (difference between OD measured on a lysate of infected cells and OD measured on a lysate of noninfected cells) after visualization with an anti-mouse IgG polyclonal antibody coupled with peroxidase (NA931V, Amersham) and

TMB supplemented with H_2O_2 (KPL) (figure 28A).

Under these conditions, the expression plasmid pcDNA-S only allows the induction of low antibody titers directed against SARS-CoV S in 3 mice out of 6 (LOG_{10}(TI)= 1.9±0.6) whereas the plasmid pcDNA-N allows the induction of anti-N antibodies at high titers (LOG_{10}(TI)= 3.9±0.3) in all the animals, and the control plasmids (pCI, pCI-HA) do not result in any detectable antibody (LOG_{10}(TI)< 1.7). The plasmid pCI-S equipped with a splice signal allows the induction of antibodies at high titers (LOG_{10}(TI)= 3.7±0.2), which are approximately 60 times higher than those observed after injection of the plasmid pcDNA-S ($p<10^{-5}$).

The efficiency of the posttranscriptional signals was studied by carrying out a dose-response study of the anti-S antibody titers induced in the BALB/c mouse as a function of the quantity of plasmid DNA used as immunogen (2 µg, 10 µg and 50 µg). This study (figure 28B) demonstrates that the posttranscriptional signal WPRE greatly improves the efficiency of gene immunization when small doses of DNA are used ($p<10^{-5}$ for a dose of 2 µg of DNA and $p<10^{-2}$ for a dose of 10 µg), whereas the effect of the CTE signal remains marginal (p=0.34 for a dose of 2 µg of DNA).

Finally, the antibodies induced in mice after gene immunization neutralize the infectivity of SARS-CoV *in vitro* (figures 29A and 29B) at titers which are consistent with the titers measured by ELISA.

In summary, the use of a splice signal and of the posttranscriptional signal WPRE of the woodchuck hepatitis virus considerably improves the induction of neutralizing antibodies directed against SARS-CoV after gene immunization with the aid of plasmid DNA directing the expression of the cDNA for SARS-CoV S.

Example 13: Diagnostic applications of the S protein

The ELISA reactivity of the recombinant Ssol polypeptide was analyzed with respect to sera from patients suffering from SARS.

The sera from probable cases of SARS tested were chosen on the basis of the results (positive or negative) of analysis of their specific reactivity toward the native antigens of SARS-CoV by immunofluorescence test on VeroE6 cells infected with SARS-CoV and/or by indirect ELISA test using as antigen a lysate of VeroE6 cells infected with SARS-CoV. The sera of these patients are identified by a serial number of the National Reference Center for Influenza Viruses and by the initials of the patient and the number of days elapsed since the onset of the symptoms. All the sera of probable cases (cf. Table XII) recognize the native antigens of SARS-CoV, with the exception of the serum 032552 of the patient VTT for whom infection with SARS-CoV could not be confirmed by RT-PCR performed on respiratory samples of days 3, 8 and 12. A panel of control sera was used as control (TV sera): they are sera collected in France before the SARS epidemic that occurred in 2003.

Table XII: Sera of probable cases of SARS

Serum	Patient	Sample collection day
031724	JYK	7
033168	JYK	38
033597	JYK	74
032632	NTM	17
032634	THA	15
032541	PHV	10
032542	NIH	17

032552	VTT	8
032633	PTU	16
032791	JLB	3
033258	JLB	27
032703	JCM	8
033153	JCM	29

Solid phases sensitized with the recombinant Ssol polypeptide were prepared by adsorption of a solution of purified Ssol polypeptide at 2 μg/ml in PBS in the wells of an ELISA plate, and then the plates are incubated overnight at 4°C and washed with PBS-Tween buffer (PBS, 0.1% Tween 20). After saturating the ELISA plates with a solution of PBS-10% skimmed milk (weight/volume) and washing in PBS-Tween, the sera to be tested (100 μl) are diluted 1/400 in PBS skimmed milk-Tween buffer (PBS, 3% skimmed milk, 0.1% Tween) and then added to the wells of the sensitized ELISA plate. The plates are incubated for 1 h at 37°C. After 3 washings with PBS-Tween buffer, the anti-human IgG conjugate labeled with peroxidase (ref. NA933V, Amersham) diluted 1/4000 in PBS-skimmed milk-Tween buffer is added, and then the plates are incubated for 1 hour at 37°C. After 6 washings with PBS-Tween buffer, the chromogen (TMB) and the substrate (H_2O_2) are added and the plates are incubated for 10 minutes protected from light. The reaction is stopped by adding a 1 N H_3PO_4 solution, and then the absorbance is measured at 450 nm with a reference at 620 nm.

The ELISA tests (figure 30) demonstrate that the recombinant Ssol polypeptide is specifically recognized by the serum antibodies of patients suffering from SARS collected at the medium or late phase of infection (≥ 10 days after the onset of the symptoms) whereas it is not significantly recognized by the serum antibodies of 2 patients (JLB and JCM) collected in the early phase of infection (3 to 8 days after the onset of the

symptoms) or by control sera of subjects not suffering from SARS. The serum antibodies of patients JLB and JCM show a seroconversion between days 3 and 27 for the first and 8 and 29 for the second after the onset of the symptoms, which confirms the specificity of the reactivity of these sera toward the Ssol polypeptide.

In conclusion, these results demonstrate that the recombinant Ssol polypeptide may be used as an antigen for the development of an ELISA test for serological diagnosis of infection with SARS-CoV.

Example 14: Vaccine applications of the recombinant soluble S protein

The immunogenicity of the recombinant Ssol polypeptide was studied in mice.

For that, a group of 6 mice was immunized at 3 weeks' interval with 10 μg of recombinant Ssol polypeptide adjuvanted with 1 mg of aluminum hydroxide (Alu-gel-S, Serva) diluted in PBS. Three successive immunizations were performed and the immune sera were collected 3 weeks after each of the immunizations (IS1, IS2, IS3). As a control, a group of mice (mock group) received aluminum hydroxide alone according to the same protocol.

The immune sera were analyzed per pool for each of the 2 groups by indirect ELISA using a lysate of VeroE6 cells infected with SARS-CoV as antigen and as a control a lysate of noninfected VeroE6 cells. The anti-SARS-CoV antibody titers are calculated as the reciprocal of the dilution producing a specific OD of 0.5 after visualization with an anti-mouse IgG(H+L) polyclonal antibody coupled with peroxidase (NA931V, Amersham) and TMB supplemented with H_2O_2 (KPL). This analysis (figure 31) shows that the immunization with

the Ssol polypeptide induces in mice, from the first immunization, antibodies directed against the native form of the SARS-CoV spicule protein present in the lysate of infected VeroE6 cells. After 2 then 3 immunizations, the anti-S antibody titers become very high.

The immune sera were analyzed per pool for each of the two groups for their capacity to seroneutralize the infectivity of SARS-CoV. 4 points of seroneutralization on FRhK-4 cells (100 TCID50 of SARS-CoV) are produced for each of the 2-fold dilutions tested from 1/20. The seroneutralizing titer is calculated according to the Reed and Munsch method as the reciprocal of the dilution neutralizing the infectivity of 2 wells out of 4. This analysis shows that the antibodies induced in mice by the Ssol polypeptide are neutralizing: the titers observed are very high after 2 and then 3 immunizations (greater than 2560 and 5120 respectively, table XIII).

Table XIII: Induction of antibodies directed against SARS-CoV after immunization with the recombinant Ssol polypeptide. The immune sera were analyzed per pool for each of the two groups for their capacity to seroneutralize the infectivity of 100 TCID50 of SARS-CoV on FRhK-4 cells. 4 points are produced for each of the 2-fold dilutions tested from 1/20. The seroneutralizing titer is calculated according to the Reed and Munsch method as the reciprocal of the dilution neutralizing the infectivity of 2 wells out of 4.

Group	**Sera**	**Neutralizing Ab**
Mock	pi	< 20
	IS1	< 20
	IS2	< 20
	IS3	< 20
Ssol	pi	< 20

	IS1	57
	IS2	> 2560
	IS3	> 5120

The neutralizing titers observed in mice immunized with the Ssol polypeptide reach levels far greater than the titers observed by Yang et al. in mice (2004, Nature, 428:561-564) and those observed by Buchholz in the hamster (2004, PNAS 101:9804-9809) which protect respectively mice and hamsters from infection with SARS-CoV. It is therefore probable that the neutralizing antibodies induced in mice after immunization with the Ssol polypeptide protect these animals against infection with SARS-CoV.

Example 15: Optimized synthetic gene for the expression in mammalian cells of the SARS-associated coronavirus (SARS-CoV) spicule (S) protein

1) Design of the synthetic gene

A synthetic gene encoding the SARS-CoV spicule protein was designed from the gene of the isolate 031589 (plasmid pSARS-S, C.N.C.M. No. I-3059) so as to allow high levels of expression in mammalian cells and in particular in cells of human origin.

For that:

- the use of codons of the wild-type gene of the isolate 031589 was modified so as to become close to the bias observed in humans and to improve the efficiency of translation of the corresponding mRNA
- the overall GC content of the gene was increased so as to extend the half-life of the corresponding mRNA
- the optionally cryptic motifs capable of interfering with an efficient expression of the

gene were deleted (splice donor and acceptor sites, polyadenylation signals, sequences very rich (> 80%) or very low (< 30%) in GC, repeat sequences, sequences involved in the formation of secondary RNA structures, TATA boxes)

- a second STOP codon was added to allow efficient termination of translation.

In addition, CpG motifs were introduced into the gene so as to increase its immunogenicity as DNA vaccine. In order to facilitate the manipulation of the synthetic gene, two BamH1 and Xho1 restriction sites were placed on either side of the open reading frame of the S protein, and the BamH1, Xho1, Nhe1, Kpn1, BspE1 and Sal1 restriction sites were avoided in the synthetic gene.

The sequence of the synthetic gene designed (gene 040530) is given in SEQ ID No: 140.

An alignment of the synthetic gene 040530 with the sequence of the wild-type gene of the isolate 031589 of SARS-CoV deposited at the C.N.C.M. under the number I-3059 (SEQ ID No: 4, plasmid pSRAS-S) is presented in figure 32.

2) Plasmid constructs

The synthetic gene SEQ ID No: 140 was assembled from synthetic oligonucleotides and cloned between the Kpn1 and Sac1 sites of the plasmid pUC-Kana in order to give the plasmid 040530pUC-Kana. The nucleotide sequence of the insert of the plasmid 040530pUC-Kana was verified by automated sequencing (Applied).

A Kpn1-Xho1 fragment containing the synthetic gene 040530 was excised from the plasmid 040530pUC-Kana and subcloned between the Nhe1 and Xho1 sites of the

expression plasmic pCI (Promega) in order to obtain the plasmid pCI-SSYNTH, deposited at the CNCM on December 1, 2004, under the number I-3333.

A synthetic gene encoding the soluble form of the S protein was then obtained by fusing the synthetic sequences encoding the ectodomain of the S protein (amino acids 1 to 1193) with those of the tag (FLAG:DYKDDDDK) via a linker BspE1 encoding the dipeptide SG. Practically, a DNA fragment encoding the ectodomain of the SARS-CoV S was amplified by PCR with the aid of the oligonucleotides 5'-ACTAGCTAGC GGATCCACCATGTTCATCTT CCTG-3' and 5'-AGTATCCGGAC TTG ATGTACT GCTCGTACTTGC-3' from the plasmid 040530pUC-Kana, digested with Nhe1 and BspE1 and then inserted between the unique Nhe1 and BspE1 sites of the plasmid pCI-Ssol, to give the plasmid pCI-SCUBE, deposited at the CNCM on December 1, 2004, under the number I-3332. The plasmids pCI-Ssol, pCI-Ssol-CTE, and pCI-Ssol-WPRE (deposited at the CNCM, on November 22, 2004, under the number I-3324) had been previously obtained by subcloning the Kpn1-Xho1 fragment excised from the plasmid pcDNA-Ssol (see technical note of DI 2004-106) between the Nhe1 and Xho1 sites of the plasmids pCI, pCI-S-CTE and pCI-S-WPRE respectively.)

The plasmids pCI-Scube and pCI-Ssol encode the same recombinant Ssol polypeptide.

3) Results

The capacity of the synthetic gene encoding the S protein to efficiently direct the expression of the SARS-CoV S in mammalian cells was compared with that of the wild-type gene after transient transfection of primate cells (VeroE6) and of human cells (293T).

In the experiment presented in figure 33 and in table XIV, monolayers of 5×10^5 VeroE6 cells or 7×10^5 293T cells in 35 mm Petri dishes were transfected with 2 μg of plasmids pCI (as control), pCI-S, pCI-S-CTE, pCI-S-WPRE and pCI-S-Ssynth and 6 μl of Fugene6 reagent according to the manufacturer's instructions (Roche). After 48 hours of incubation at 37°C and under 5% CO_2, cell extracts were prepared in loading buffer according to Laemmli, separated on 8% SDS polyacrylamide gel and then transferred onto a PVDF membrane (BioRad). The detection of this immunoblot (Western blot) was carried out with the aid of an anti-S rabbit polyclonal serum (immune serum of the rabbit P11135: cf example 4 above) and of donkey polyclonal antibodies directed against rabbit IgGs and coupled with peroxidase (NA934V, Amersham). The immunoblot was quantitatively visualized by luminescence with the aid of the ECL+ kit (Amersham) and acquisition on a digital imaging device (FluorS, BioRad).

The analysis of the results obtained with the software QuantityOne v4.2.3 (BioRad) shows that in this experiment, the plasmid pCI-Synth allows the transient expression of the S protein at high levels in the VeroE6 and 293T cells, whereas the plasmid pCI-S does not make it possible to induce expression at sufficient levels to be detected. The expression levels observed are of the order of twice as high as those observed with the plasmid pCI-S-WPRE.

Table XIV: Use of a synthetic gene for the expression of the SARS-CoV S. Cell extracts prepared 48 hours after transfection of VeroE6 or 293T cells with the plasmids pCI, pCI-S, pCI-S-CTE, pCI-S-WPRE and pCI-S-Ssynth were separated on 8% SDS acrylamide gel and analyzed by Western blotting with the aid of an anti-S rabbit polyclonal antibody and an anti-rabbit IgG(H+L) polyclonal antibody coupled with peroxidase (NA934V,

Amersham). The Western blot is visualized by luminescence (ECL+, Amersham) and acquisition on a digital imaging device (FluorS, BioRad). The expression levels of the S protein were measured by quantifying the two predominant bands identified on the image (see figure 33) and are indicated according to an arbitrary scale where the value 1 represents the level measured after transfection of the plasmid pCI-S-WPRE.

Plasmid	**VeroE6**	**293T**
pCI	0.0	0.0
pCI-S	≤ 0.1	≤ 0.1
pCI-S-CTE	0.5	≤ 0.1
pCI-S-WPRE	1.0	1.0
pCI-Ssynth	1.8	1.9

In a second instance, the capacity of the synthetic gene Scube to efficiently direct the synthesis and the secretion of the Ssol polypeptide by mammalian cells was compared with that of the wild-type gene after transient transfection of hamster cells (BHK-21) and of human cells (293T).

In the experiment presented in table XV, monolayers of 6×10^5 BHK-21 cells and 7×10^5 293T cells in 35 mm Petri dishes were transfected with 2 μg of plasmids pCI (as control), pCI-Ssol, pCI-Ssol-CTE, pCI-Ssol-WPRE and pCI-Scube and 6 μl of Fugene6 reagent according to the manufacturer's instructions (Roche). After 48 hours of incubation at 37°C and under 5% CO_2, the cellular supernatants were collected and quantitatively analyzed for the secretion of the Ssol polypeptide by a capture ELISA test specific for the Ssol polypeptide.

Analysis of the results shows that, in this experiment, the plasmid pCI-Scube allows the expression of the Ssol polypeptide at levels 8 times (BHK-21 cells) to 20 times (293T cells) higher than the plasmid pCI-Ssol.

The levels of expression observed are of the order of twice (293T cells) to 5 times (BHK-21 cells) as high as those observed with the plasmid pCI-Ssol-WPRE.

Table XV: Use of a synthetic gene for the expression of the Ssol polypeptide. The supernatants were harvested 48 hours after transfection of BHK or 293T cells with the plasmids pCI, pCI-Ssol, pCI-Ssol-CTE, pCI-Ssol-WPRE and pCI-Scube and quantitatively analyzed for the secretion of the Ssol polypeptide by an ELISA test specific for the Ssol polypeptide. The transfections were carried out in duplicate and the results are presented in the form of means and standard deviations of the concentrations of Ssol polypeptide (ng/ml) measured in the supernatants.

Plasmid	BHK	293T
pci	< 20	< 20
pCI-Ssol	< 20	56 ± 10
pCI-Ssol-CTE	< 20	63 ± 8
pCI-Ssol-WPRE	28 ± 1	531 ± 15
pCI-Scube	152 ± 6	1140 ± 20

In summary, these results show that the expression, in mammalian cells, of the synthetic gene 040530 encoding SARS-CoV S under the control of RNA polymerase II promoter sequences is much more efficient than that of the wild-type gene of the 031589 isolate. This expression is even more efficient than that directed by the wild-type gene in the presence of the WPRE sequences of the woodchuck hepatitis virus.

4) Applications

The use of the synthetic gene 040530 encoding SARS-CoV S or its Scube variant encoding the polypeptide Ssol is capable of advantageously replacing the wild-type gene

in numerous applications where the expression of S is necessary at high levels. In particular in order to:

- improve the efficiency of gene immunization with plasmids of the pCI-Ssynth or even pCI-Ssynth-CTE or pCI-Ssynth-WPRE type
- establish novel cell lines expressing higher quantities of the S protein or of the Ssol polypeptide with the aid of recombinant lentiviral vectors carrying the Ssynth gene or the Scube gene respectively
- improve the immunogenicity of the recombinant lentiviral vectors allowing the expression of the S protein or of the Ssol polypeptide
- improve the immunogenicity of live vectors allowing the expression of the S protein or of the Ssol polypeptide like recombinant vaccinia viruses or recombinant measles viruses (see examples 16 and 17 below)

<u>Example 16</u>: Expression of the SARS-associated coronavirus (SARS-CoV) spicule (S) protein with the aid of recombinant vaccinia viruses

Vaccine application

Application to the production of a soluble form of the spicule (S) protein and design of a serological test for SARS

1) Introduction

The aim of this example is to evaluate the capacity of recombinant vaccinia viruses (VV) expressing various SARS-associated coronavirus (SARS-CoV) antigens to constitute novel vaccine candidates against SARS and a means of producing recombinant antigens in mammalian cells.

For that, the inventors focused on the SARS-CoV spicule (S) protein which makes it possible to induce, after gene immunization in animals, antibodies neutralizing

the infectivity of SARS-CoV, and a soluble and secreted form of this protein, the Ssol polypeptide, which is composed of the ectodomain (aa 1-1193) of S fused at its C-ter end with a tag FLAG (DYKDDDDK) via a BspE1 linker encoding the SG dipeptide. This Ssol polypeptide exhibits an antigenicity similar to that of the S protein and allows, after injection into mice in the form of a purified protein adjuvanted with aluminum hydroxide, the induction of high neutralizing antibody titers against SARS-CoV.

The various forms of the S gene were placed under the control of the promoter of the 7.5K gene and then introduced into the thymidine kinase (TK) locus of the Copenhagen strain of the vaccinia virus by double homologous recombination *in vivo*. In order to improve the immunogenicity of the recombinant vaccinia viruses, a synthetic late promoter was chosen in place of the 7.5K promoter, in order to increase the production of S and Ssol during the late phases of the viral cycle.

After having isolated the recombinant vaccinia viruses and verified their capacity to express the SARS-CoV S antigen, their capacity to induce in mice an immune response against SARS was tested. After having purified the Ssol antigen from the supernatant of infected cells, an ELISA test for serodiagnosis of SARS was designed, and its efficiency was evaluated with the aid of sera from probable cases of SARS.

2) Construction of the recombinant viruses

Recombinant vaccinia viruses directing the expression of the S glycoprotein of the 031589 isolate of SARS-CoV and of a soluble and secreted form of this protein, the Ssol polypeptide, under the control of the 7.5K promoter were obtained. With the aim of increasing the levels of expression of S and Ssol, recombinant viruses

in which the cDNAs for S and for Ssol are placed under the control of a late synthetic promoter were also obtained.

The plasmid pTG186poly is a transfer plasmid for the construction of recombinant vaccinia viruses (Kieny, 1986, Biotechnology, 4:790-795). As such, it contains the VV thymidine kinase gene into which the promoter of the 7.5K gene has been inserted followed by a multiple cloning site allowing the insertion of heterologous genes (figure 34A). The promoter of the 7.5K gene in fact contains a tandem of two promoter sequences that are respectively active during the early (P_E) and late (P_L) phases of the vaccinia virus replication cycle. The BamH1-Xho1 fragments were excised from the plasmids pTRIP-S and pcDNA-Ssol respectively and inserted between the BamH1 and Sma1 sites of the plasmid pTG186poly in order to give the plasmids pTG-S and pTG-Ssol (figure 34A). The plasmids pTG-S and pTG-Ssol were deposited at the CNCM, on December 2, 2004, under the numbers I-3338 and I-3339, respectively.

The plasmids pTN480, pTN-S and pTN-Ssol were obtained from the plasmids pTG186poly, pTG-S and pTG-Ssol respectively, by substituting the Nde1-Pst1 fragment containing the 7.5K promoter by a DNA fragment containing the synthetic late promoter 480, which was obtained by hybridization of the oligonucleotides 5'-TATGAGCTTT TTTTTTTTTT TTTTTTTGGC ATATAAATAG ACTCGGCGCG CCATCTGCA-3' and 5'- GATGGCGCGCCGAGTCTATT TATATGCCAA AAAAAAAAAA AAAAAAAAGC TCA-3' (figure 34B). The insert was sequenced with the aid of a BigDye Terminator v1.1 kit (Applied Biosystems) and an automated sequencer ABI377. The sequence of the late synthetic promoter 480 as cloned into the transfer plasmids of the pTN series is indicated in figure 34C. The plasmids pTN-S and pTN-Ssol were deposited at the CNCM, on December 2, 2004, under the numbers I-3340 and I-3341, respectively.

The recombinant vaccinia viruses were obtained by double homologous recombination *in vivo* between the TK cassette of the transfer plasmids of the series pTG and pTN and the TK gene of the Copenhagen strain of the vaccinia virus according to a procedure described by Kieny et al. (1984, Nature, 312:163-166). Briefly, CV-1 cells are transfected with the aid of DOTAP (Roche) with genomic DNA of the Copenhagen strain of the vaccinia virus and each of the transfer plasmids of the pTG and pTN series described above, and then superinfected with the helper vaccinia virus VV-ts7 for 24 hours at 33°C. The helper virus is counter-selected by incubation at 40°C for 2 days and then the recombinant viruses (TK- phenotype) selected by two cloning cycles under agar medium on 143Btk- cells in the presence of BuDr (25 µg/ml). The 6 viruses VV-TG, VV-TG-S, VV-TG-Ssol, VV-TN, VV-TN-S, and VV-TN-Ssol are respectively obtained with the aid of the transfer plasmids pTG186poly, pTG-S, pTG-Ssol, pTN480, pTN-S, pTN-Ssol. The viruses VV-TG and VV-TN do not express any heterologous gene and were used as TK- control in the experiments. The preparations of recombinant viruses were performed on monolayers of CV-1 or BHK-21 cells and the titer in plaque forming units (p.f.u) determined on CV-1 cells according to Earl and Moss (1998, Current Protocols in Molecular Biology, 16.16.1-16.16.13).

3) Characterization of the recombinant viruses

The expression of the transgenes encoding the S protein and the Ssol polypeptide was assessed by Western blotting.

Monolayers of CV-1 cells were infected at a multiplicity of 2 with various recombinant vaccinia viruses VV-TG, VV-TG-S, VV-TG-Ssol, VV-TN, VV-TN-S and

VV-TN-Ssol. After 18 hours of incubation at 37°C and under 5% CO2, cellular extracts were prepared in loading buffer according to Laemmli, separated on 8% SDS polyacrylamide gel and then transferred onto a PVDF membrane (BioRad). The detection of this immunoblot (Western blot) was performed with the aid of an anti-S rabbit polyclonal serum (immune serum from the rabbit P11135: cf. example 4) and donkey polyclonal antibodies directed against rabbit IgGs and coupled with peroxidase (NA934V, Amersham). The bound antibodies were visualized by luminescence with the aid of the ECL+ kit (Amersham) and autoradiography films Hyperfilm MP (Amersham).

As shown in figure 35A, the recombinant virus VV-TN-S directs the expression of the S protein at levels which are comparable to those which can be observed 8 h after infection with SARS-CoV but which are much higher than those which can be observed after infection with VV-TG-S. In a second experiment (figure 35B), the analysis of variable quantities of cellular extracts shows that the levels of expression observed after infection with viruses of the TN series (VV-TN-S and VV-TN-Ssol) are about 10 times as high as those observed with the viruses of the TG series (VV-TG-S and VV-TG-Ssol, respectively). In addition, the Ssol polypeptide is secreted into the supernatant of CV-1 cells infected with the VV-TN-Ssol virus more efficiently than in the supernatant of cells infected with VV-TG-Ssol (figure 36A). In this experiment, the VV-TN-Sflag virus was used as a control because it expresses the membrane form of the S protein fused at its C-ter end with the FLAG tag. The Sflag protein is not detected in the supernatant of cells infected with VV-TN-Sflag, demonstrating that the Ssol polypeptide is indeed actively secreted after infection with VV-TN-Ssol.

These results demonstrate that the recombinant vaccinia viruses are indeed carriers of the transgenes and allow the expression of the SRAS glycoprotein in its membrane form (S) or in a soluble or secreted form (Ssol). The vaccinia viruses carrying the synthetic promoter 480 allow the expression of S and the secretion of Ssol at levels much higher than the viruses carrying the promoter of the 7.5K gene.

4) Application to the production of a soluble form of SARS-CoV S. Purification of this recombinant antigen and diagnostic applications

The BHK-21 line is the cell line which secretes the highest quantities of Ssol polypeptide after infection with the VV-TN-Ssol virus among the lines tested (BHK-21, CV1, 293T and FrhK-4, figure 36B); it allows the quantitative production and purification of the recombinant Ssol polypeptide. In a typical experiment where the experimental conditions for infection, production and purification were optimized, the BHK-21 cells are inoculated in standard culture medium (pyruvate-free DMEM containing 4.5 g/l of glucose and supplemented with 5% TPB, 5% FCS, 100 U/ml of penicillin and 100 µg/ml of streptomycin) in the form of a subconfluent monolayer (10 million cells for each 100 cm^2 in 25 ml of medium). After 24 h of incubation at 37°C under 5% CO_2, the cells are infected at an M.O.I. of 0.03 and the standard medium replaced with the secretion medium where the quantity of FCS is reduced to 0.5% and the TPB eliminated. The culture supernatant is removed after 2.5 days of incubation at 35°C and under 5% CO_2 and the vaccinia virus inactivated by addition of Triton X-100 (0.1%). After filtration on 0.1 µm polyethersulfone (PES) membrane, the recombinant Ssol polypeptide is purified by affinity chromatography on an anti-FLAG matrix with elution with a solution of FLAG peptide (DYKDDDDK) at

100 µg/ml in TBS (50 mM Tris, pH 7.4, 150 mM NaCl).

The analysis by 8% SDS acrylamide gel stained with silver nitrate identified a predominant polypeptide whose molecular mass is about 180 kD and whose degree of purity is greater than 90% (figure 37). The concentration of the purified Ssol recombinant polypeptide was determined by comparison with molecular mass markers and estimated at 24 ng/µl.

This purified Ssol polypeptide preparation makes it possible to produce a calibration series in order to measure, with the aid of a capture ELISA test, the Ssol concentrations present in the culture supernatants. According to this test, the BHK-21 line secretes about 1 µg/ml of Ssol polypeptide under the production conditions described above. In addition, the purification scheme presented makes it possible to purify of the order of 160 µg of Ssol polypeptide per liter of culture supernatant.

The ELISA reactivity of the recombinant Ssol polypeptide was analyzed toward sera from patients suffering from SARS.

The sera of probable cases of SARS tested were chosen on the basis of the results (positive or negative) of analysis of their specific reactivity toward the native antigens of SARS-CoV by immunofluorescence test on VeroE6 cells infected with SARS-CoV and/or by indirect ELISA test using, as antigen, a lysate of VeroE6 cells infected with SARS-CoV. The sera of these patients are identified by a serial number of the National Reference Center for Influenza Viruses and by the patient's initials and the number of days elapsed since the onset of the symptoms. All the sera of probable cases (cf. table XVI) recognize the native antigens of SARS-CoV with the exception of the serum 032552 of the patient

VTT, for which infection with SARS-CoV could not be confirmed by RT-PCR performed on respiratory samples of days 3, 8 and 12. A panel of control sera was used as control (TV sera): they are sera collected in France before the SARS epidemic which occurred in 2003.

Table XVI: Sera of probable cases of SARS

Serum	Patient	Sample collection day
033168	JYK	38
033597	JYK	74
032632	NTM	17
032634	THA	15
032541	PHV	10
032542	NIH	17
032552	VTT	8
032633	PTU	16

Solid phases sensitized with the recombinant Ssol polypeptide were prepared by adsorption of a solution of purified Ssol polypeptide at 4 µg/ml in PBS in the wells of an ELISA plate. The plates are incubated overnight at 4°C and then washed with PBS-Tween buffer (PBS, 0.1% Tween 20). After washing with PBS-Tween, the sera to be tested (100 µl) are diluted 1/100 and 1/400 in PBS-skimmed milk-Tween buffer (PBS, 3% skimmed milk, 0.1% Tween) and then added to the wells of the sensitized ELISA plate. The plates are then incubated for 1 h at 37°C. After 3 washings with PBS-Tween buffer, the anti-human IgG conjugate labeled with peroxidase (ref. NA933V, Amersham) diluted 1/4000 in PBS-skimmed milk-Tween buffer is added and then the plates are incubated for one hour at 37°C. After 6 washings with PBS-Tween buffer, the chromogen (TMB) and the substrate (H_2O_2) are added and the plates are incubated for 10 minutes protected from light. The

reaction is stopped by adding a 1M solution of H_3PO_4 and then the absorbance is measured at 450 nm with a reference at 620 nm.

The ELISA tests (figure 38) demonstrate that the recombinant Ssol polypeptide is specifically recognized by the serum antibodies of patients suffering from SARS, collected at the middle or late phase of infection (≥ 10 days after the onset of the symptoms), whereas it is not significantly recognized by the serum antibodies of the control sera of subjects not suffering from SARS.

In conclusion, these results demonstrate that the recombinant Ssol polypeptide can be purified from the supernatant of mammalian cells infected with the recombinant vaccinia virus VV-TN-Ssol and can be used as antigen for developing an ELISA test for serological diagnosis of infection with SARS-CoV.

5. Vaccine applications

The immunogenicity of the recombinant vaccinia viruses was studied in mice.

For that, groups of 7 BALB/c mice were immunized by the i.v. route twice at 4 weeks' interval with 10^6 p.f.u. of recombinant vaccinia viruses VV-TG, VV-TG-S, VV-TG-Ssol, VV-TN, VV-TN-S and VV-TN-Ssol and, as a control, VV-TG-HA which directs the expression of hemagglutinin of the A/PR/8/34 strain of the influenza virus. The immune sera were collected 3 weeks after each of the immunizations (IS1, IS2).

The immune sera were analyzed per pool for each of the groups by indirect ELISA using a lysate of VeroE6 cells infected with SARS-CoV as antigen and, as control, a lysate of noninfected VeroE6 cells. The anti-SARS-CoV

antibody titers (TI) are calculated as the reciprocal of the dilution producing a specific OD of 0.5 after visualization with an anti-mouse IgG(H+L) polyclonal antibody coupled with peroxidase (NA931V, Amersham) and TMB supplemented with H_2O_2 (KPL). This analysis (figure 39A) shows that immunization with the virus VV-TG-S and VV-TN-S induces in mice, from the first immunization, antibodies directed against the native form of the SARS-CoV spicule protein present in the lysate of infected VeroE6 cells. The responses induced by the VV-TN-S virus are higher than those induced by the VV-TG-S virus after the first (TI = 740 and TI = 270 respectively) and the second (TI = 3230 and TI = 600 respectively) immunization. The VV-TN-Ssol virus induces high anti-SARS-CoV antibody titers after two immunizations (TI = 640), whereas the virus VV-TG-Ssol induces a response at the detection limit (TI = 40).

The immune sera were analyzed per pool for each of the groups for their capacity to seroneutralize the infectivity of SARS-CoV. 4 seroneutralization points on FRhK-4 cells (100 TCID50 of SARS-CoV) are produced for each of the 2-fold dilutions tested from 1/20. The seroneutralizing titer is calculated according to the Reed and Munsch method as the reciprocal of the dilution neutralizing the infectivity of 2 wells out of 4. This analysis shows that the antibodies induced in mice by the vaccinia viruses expressing the S protein or the Ssol polypeptide are neutralizing and that the viruses with synthetic promoters are more efficient immunogens than the viruses carrying the 7.5K promoter: the highest titers (640) are observed after 2 immunizations with the virus VV-TN-S (figure 39B).

The protective power of the neutralizing antibodies induced in mice after immunization with the recombinant vaccinia viruses is evaluated with the aid of a challenge infection with SARS-CoV.

6) Other applications

Third generation recombinant vaccinia viruses are constructed by substituting the wild-type sequences of the S and Ssol genes by synthetic genes optimized for the expression in mammalian cells, described above. These recombinant vaccinia viruses are capable of expressing larger quantities of S and Ssol antigens and therefore of exhibiting increased immunogenicity.

The recombinant vaccinia virus VV-TN-Ssol can be used for the quantitative production and purification of the Ssol antigen for diagnostic (serology by ELISA) and vaccine (subunit vaccine) applications.

Example 17: Recombinant measles virus expressing the SARS-associated coronavirus (SARS-CoV) spicule (S) protein. Vaccine applications.

1) Introduction

The measles vaccine (MV) induces a lasting protective immunity in humans after a single injection (Hilleman, 2002, Vaccine, 20: 651-665). The protection conferred is very robust and is based on the induction of an antibody response and of a CD4 and CD8 cell response. The MV genome is very stable and no reversion of the vaccine strains to virulence has ever been observed. The measles virus belongs to the genus *Morbillivirus* of the *Paramyxoviridae* family; it is an enveloped virus whose genome is a 16 kb single-stranded RNA of negative polarity (figure 40A) and whose exclusively cytoplasmic replication cycle excludes any possibility of integration into the genome of the host. The measles vaccine is thus one of the most effective and one of the safest live vaccines used in the human population. Frédéric Tangy's team recently developed an expression

vector on the basis of the Schwarz strain of the measles virus, which is the safest attenuated strain and the most widely used in humans as vaccine against measles. This vaccine strain may be isolated from an infectious molecular clone while preserving its immunogenicity in primates and in mice that are sensitive to the infection. It constitutes, after insertion of additional transcription units, a vector for the expression of heterologous sequences (Combredet, 2003, J. Virol. 77: 11546-11554). In addition, a recombinant MV Schwarz expressing the envelope glycoprotein of the West Nile virus (WNV) induces an effective and lasting antibody response which protects mice from a lethal challenge infection with WNV (Despres et al., 2004, J. Infect. Dis., in press). All these characteristics make the attenuated Schwarz strain of the measles virus an extremely promising candidate vector for the construction of novel recombinant live vaccines.

The aim of this example is to evaluate the capacity of recombinant measles viruses (MV) expressing various SARS-associated coronavirus (SARS-CoV) antigens to constitute novel candidate vaccines against SARS.

The inventors focused on the SARS-CoV spicule (S) protein, which makes it possible to induce, after gene immunization in animals, antibodies neutralizing the infectivity of SARS-CoV, and on a soluble and secreted form of this protein, the Ssol polypeptide, which is composed of the ectodomain (aa 1-1193) of S fused at its C-ter end with a FLAG tag (DYKDDDDK) via a BspE1 linker encoding the SG dipeptide. This Ssol polypeptide exhibits a similar antigenicity to that of the S protein and allows, after injection into mice in the form of a purified protein adjuvanted with aluminum hydroxide, the induction of high neutralizing antibody titers against SARS-CoV.

The various forms of the S gene were introduced in the form of an additional transcription unit between the P (phosphoprotein) and M (matrix) genes into the cDNA of the Schwarz strain of MV previously described (Combredet, 2003, J. Virol. 77: 11546-11554; EP application No. 02291551.6 of June 20, 2002, and EP application No. 02291550.8 of June 20, 2002). After having isolated the recombinant viruses MVSchw2-SARS-S and MVSchw2-SARS-Ssol and checked their capacity to express the SARS-CoV S antigen, their capacity to induce a protective immune response against SARS in mice and then in monkeys was tested.

2) Construction of the recombinant viruses

The plasmid pTM-MVSchw-ATU2 (figure 40B) contains an infectious cDNA corresponding to the antigenome of the Schwarz vaccine strain of the measles virus (MV) into which an additional transcription unit (ATU) has been introduced between the P (phosphoprotein) and M (matrix) genes (Combredet, 2003, Journal of Virology, 77: 11546-11554). Recombinant genomes MVSchw2-SARS-S and MVSchw2-SARS-Ssol of the measles virus were constructed by inserting ORFs of the S protein and of the Ssol polypeptide into the additional transcription unit of the MVSchw-ATU2 vector.

For that, a DNA fragment containing the SARS-CoV S cDNA was amplified by PCR with the aid of the oligonucleotides 5'-ATACGTACGA CCATGTTTAT TTTCTTATTA TTTCTTACTC TCACT-3' and 5'-ATAGCGCGCT CATTATGTGT AATGTAATTT GACACCCTTG-3' using the plasmid pcDNA-S as template and then inserted into the plamid pCR®2.1-TOPO (Invitrogen) in order to obtain the plasmid pTOPO-S-MV. The two oligonucleotides used contain restriction sites BsiW1 and BssHII, so as to allow subsequent insertion into the measles vector, and were designed so as to generate a sequence of 3774 nt including the codons for

initiation and termination, so as to observe the rule of 6 which stipulates that the length of the genome of a measles virus must be divisible by 6 (Calain & Roux, 1993, J. Virol., 67: 4822-4830; Schneider et al., 1997, Virology, 227: 314-322). The insert was sequenced with the aid of a BigDye Terminator v1.1 kit (Applied Biosystems) and an automated sequencer ABI377.

To express a soluble and secreted form of SARS-CoV S, a plasmid containing the cDNA of the Ssol polypeptide corresponding to the ectodomain (aa 1-1193) of SARS-CoV S fused at its C-ter end with the sequence of a FLAG tag (DYKDDDDK) via a BspE1 linker encoding the SG dipeptide was then obtained. For that, a DNA fragment was amplified with the aid of the oligonucleotides 5'-CCATTTCAAC AATTTGGCCG-3' and 5'-ATAGGATCCG CGCGCTCATT ATTTATCGTC GTCATCTTTA TAATC-3' from the plasmid pCDNA-Ssol and then inserted into the plasmid pTOPO-S-MV between the Sal1 and BamH1 sites in order to obtain the plasmid pTOPO-S-MV-SF. The sequence generated is 3618 nt long between the BsiW1 and BssHII sites and observes the rule of 6. The insert was sequenced as indicated above.

The BsiW1-BssHII fragments containing the cDNAs for the S protein and the Ssol polypeptide were then excised by digestion of the plasmids pTOPO-S-MV and pTOPO-S-MV-SF and then subcloned between the corresponding sites of the plasmid pTM-MVSchw-ATU2 in order to give the plasmids pTM-MVSchw2-SARS-S and pTM-MVSchw2-SARS-Ssol (figure 40B). These two plasmids were deposited at the C.N.C.M. on December 1, 2004, under the numbers I-3326 and I-3327, respectively.

The recombinant measles viruses corresponding to the plasmids pTM-MVSchw2-SARS-S and pTM-MVSchw2-SARS-Ssol were obtained by reverse genetics according to the system based on the use of a helper cell line,

described by Radecke et al. (1995, Embo J., 14: 5773-5784) and modified by Parks et al. (1999, J. Virol., 73: 3560-3566). Briefly, the helper cells 293-3-46 are transfected according to the calcium phosphate method with 5 μg of the plasmids pTM-MVSchw2-SARS-S or pTM-MVSchw2-SARS-Ssol and 0.02 μg of the plasmid pEMC-La directing the expression of the MV L polymerase (gift from M.A. Billeter). After incubating overnight at 37°C, a heat shock is produced for 2 hours at 43°C and the transfected cells are transferred onto a monolayer of Vero cells. For each of the two plasmids, syncytia appeared after 2 to 3 days of coculture and were transferred successively onto monolayers of Vero cells at 70% confluence in 35 mm Petri dishes and then in 25 and 75 cm^2 flasks. When the syncytia have reached 80-90% confluence, the cells are recovered with the aid of a scraper and then frozen and thawed once. After low-speed centrifugation, the supernatant containing the virus is stored in aliquots at -80°C. The titers of the recombinant viruses MVSchw2-SARS-S and MVSchw2-SARS-Ssol were determined by limiting dilution on Vero cells and the titer as dose infecting 50% of the wells ($TCID_{50}$) calculated according to the Kärber method.

3) Characterization of the recombinant viruses

The expression of the transgenes encoding the S protein and the Ssol polypeptide was assessed by Western blotting and immunofluorescence.

Monolayers of Vero cells in T-25 flasks were infected at a multiplicity of 0.05 by various passages of the two viruses MVSchw2-SARS-S and MVSchw2-SARS-Ssol and the wild-type virus MWSchw as a control. When the syncytia had reached 80 to 90% confluence, cytoplasmic extracts were prepared in an extraction buffer (150 mM NaCl, 50 mM Tris-HCl, pH 7.2, 1% Triton X-100, 0.1% SDS, 1% DOC) and then diluted in loading buffer

according to Laemmli, separated on 8% SDS polyacrylamide gel and transferred onto a PVDF membrane (BioRad). The detection of this immunoblot (Western blot) was carried out with the aid of an anti-S rabbit polyclonal serum (immune serum of the rabbit P11135: cf. example 4 above) and donkey polyclonal antibodies directed against rabbit IgGs and coupled with peroxidase (NA934V, Amersham). The bound antibodies were visualized by luminescence with the aid of the ECL+ kit (Amersham) and Hyperfilm MP autoradiography films (Amersham).

Vero cells in monolayers on glass slides were infected with the two viruses MVSchw2-SARS-S and MVSchw2-SARS-Ssol and the wild-type virus MWSchw as a control at multiplicities of infection of 0.05. When the syncytia had reached 90 to 100% (MVSchw2-SARS-Ssol virus) or 30 to 40% (MVSchw2-SARS-S, MWSchw) confluence, the cells were fixed in a 4% PBS-PFA solution, permeabilized with a PBS solution containing 0.2% Triton and then labeled with rabbit polyclonal antibodies hyperimmunized with purified and inactivated SARS-CoV virions and with an anti-rabbit IgG(H+L) goat antibody conjugate coupled with FITC (Jackson).

As shown in figures 41 and 42, the recombinant viruses MVSchw2-SARS-S and MVSchw2-SARS-Ssol direct the expression of the S protein and the Ssol polypeptide respectively at levels comparable to those which can be observed 8 h after infection with SARS-CoV. The expression of these polypeptides is stable after 3 passages of the recombinant viruses in cell culture. These results demonstrate that the recombinant measles viruses are indeed carriers of the transgenes and allow the expression of the SARS glycoprotein in its membrane form (S) or in a soluble form (Ssol). The Ssol polypeptide is expected to be secreted by cells infected with the MVSchw2-SARS-Ssol virus as is the

case when this same polypeptide is expressed in mammalian cells after transient transfection of the corresponding sequences (cf. example 11 above).

4) Applications

Having shown that the viruses MVSchw2-SARS-S and MVSchw2-SARS-Ssol allow the expression of the SARS-CoV S, their capacity to induce a protective immune response against SARS-CoV in $CD46^{+/-}$ IFN-$\alpha\beta R^{-/-}$ mice, which is sensitive to infection by MV, is evaluated. The antibody response of the immunized mice is evaluated by ELISA test against the native antigens of SARS-CoV and for their capacity to neutralize the infectivity of SARS-CoV *in vitro*, using the methodologies described above. The protective power of the response will be evaluated by measuring the reduction in the pulmonary viral load 2 days after a nonlethal challenge infection with SARS-CoV.

Second generation recombinant measles viruses are constructed by substituting the wild-type sequences of the S and Sol genes by synthetic genes optimized for expression in mammalian cells, described in example 15 above. These recombinant measles viruses are capable of expressing larger quantities of the S and Ssol antigens and therefore of exhibiting increased immunogenicity.

Alternatively, the wild-type or synthetic genes encoding the S protein or the Ssol polypeptide may be inserted into the measles vector MVSchw-ATU3 in the form of an additional transcription unit located between the H and L genes, and then the recombinant viruses produced and characterized in a similar manner. This insertion is capable of generating recombinant viruses possessing different characteristics (multiplication of the virus, level of expression of the transgene) and possibly an improved immunogenicity

compared with those obtained after insertion of the transgenes between the P and N genes.

The recombinant measles virus MVSchw2-SARS-Ssol may be used for the quantitative production and the purification of the Ssol antigen for diagnostic and vaccine applications.

Example 18: Other applications linked to the S protein

a) The lentiviral vectors allowing the expression of S or Ssol (or even of fragments of S) can constitute a recombinant vaccine against SARS-CoV, to be used in human or veterinary prophylaxis. In order to demonstrate the feasibility of such a vaccine, the immunogenicity of the recombinant lentiviral vectors TRIP-SD/SA-S-WPRE and TRIP-SD/SA-Ssol-WPRE is studied in mice.

b) Monoclonal antibodies are produced with the aid of the recombinant Ssol polypeptide. According to the results presented in example 14 above, these antibodies or at least the majority of them will recognize the native form of the SARS-CoV S and will be capable of diagnostic and/or prophylactic applications.

c) A serological test for SARS is developed with the Ssol polypeptide used as antigen and the double epitope methodology.